Journal of Experimental Psychology: General
2016, Vol. 145, No. 9, 1206–1227

0096-3445/16/$12.00 http://dx.doi.org/10.1037/xge0000182

Separate But Correlated: The Latent Structure of Space and Mathematics Across Development

Kelly S. Mix
Michigan State University

Susan C. Levine
University of Chicago

Yi-Ling Cheng
Michigan State University

Chris Young
University of Chicago

D. Zachary Hambrick
Michigan State University

Raedy Ping
University of Chicago

Spyros Konstantopoulos
Michigan State University

The relations among various spatial and mathematics skills were assessed in a cross-sectional study of 854 children from kindergarten, third, and sixth grades (i.e., 5 to 13 years of age). Children completed a battery of spatial mathematics tests and their scores were submitted to exploratory factor analyses both within and across domains. In the within domain analyses, all of the measures formed single factors at each age, suggesting consistent, unitary structures across this age range. Yet, as in previous work, the 2 domains were highly correlated, both in terms of overall composite score and pairwise comparisons of individual tasks. When both spatial and mathematics scores were submitted to the same factor analysis, the 2 domain specific factors again emerged, but there also were significant cross-domain factor loadings that varied with age. Multivariate regressions replicated the factor analysis and further revealed that mental rotation was the best predictor of mathematical performance in kindergarten, and visual-spatial working memory was the best predictor of mathematical performance in sixth grade. The mathematical tasks that predicted the most variance in spatial skill were place value (K, 3rd, 6th), word problems (3rd, 6th), calculation (K), fraction concepts (3rd), and algebra (6th). Thus, although spatial skill and mathematics each have strong internal structures, they also share significant overlap, and have particularly strong cross-domain relations for certain tasks.

Keywords: spatial cognition, mathematics, cognitive development, elementary age

Supplemental materials: http://dx.doi.org/10.1037/xge0000182.supp

Children and adults who perform better on spatial tasks also perform better on tests of mathematical ability (Burnett, Lane, & Dratt, 1979; Casey, Nuttall, & Pezaris, 2001; Casey, Nuttall, Pezaris, & Benbow, 1995; Delgado & Prieto, 2004; Geary, Saults, Liu, & Hoard, 2000; Lubinski & Benbow, 1992; Robinson, Abbott, Berninger, & Busse, 1996). Some theorists have argued that the two are related because mathematics, along with other complex concepts, is mentally represented in a spatial format (Barsalou, 2008; Lakoff & Nunez, 2000). Others have shown that similar neural circuits are activated when people process spatial and numerical information (Hubbard, Piazza, Pinel, & Dehaene, 2005; Walsh, 2003), and there is behavioral evidence the two are connected (e.g., Dehaene, Bossini, & Giraux,

Editor's Note. Daniel Ansari served as the action editor for this article.—IG

Kelly S. Mix, Department of Counseling, Educational Psychology, and Special Education, Michigan State University; Susan C. Levine, Department of Psychology, University of Chicago; Yi-Ling Cheng, Department of Counseling, Educational Psychology, and Special Education, Michigan State University; Chris Young, Department of Psychology, University of Chicago; D. Zachary Hambrick, Department of Psychology, Michigan State University; Raedy Ping, Department of Psychology, University of Chicago; Spyros Konstantopoulos, Department of Counseling, Educational Psychology, and Special Education, Michigan State University.

We thank all the children who participated in this study, as well as their parents and school district personnel, for their generous cooperation. Special thanks go to Mark Reckase for his excellent statistical advice. This research was supported by Institute of Education Sciences Grant R305A120416 to Kelly S. Mix and Susan C. Levine, as well as National Science Foundation Spatial Intelligence Learning Center Grants SBE-1041707 and SBE-0541957 to Susan C. Levine. The opinions and positions expressed in this report are the authors' and do not necessarily represent the opinions and positions of the Institute of Education Sciences or the U.S. Department of Education.

Correspondence concerning this article should be addressed to Kelly S. Mix, who is now at the Department of Human Development and Quantitative Methods, College of Education, University of Maryland, 3304 Benjamin Building, College Park, MD 20742. E-mail: kmix@umd.edu

1993; McKenzie, Bull, & Gray, 2003). Thus there is reason to think the relation is based on shared processing.

But what shared processing might this be? Both domains are comprised of multiple skills and concepts, so when we say space and mathematics are related, what kind of spatial skill do we mean? What kind of mathematics? Indeed, several studies have reported that the relations between spatial skill and mathematics shift depending on the demands of particular tasks (Caviola, Mammarella, Cornoldi, & Lucangeli, 2012; Robert & LeFevre, 2013; Trbovich & LeFevre, 2003), suggesting these cross-domain connections may be more specific than is currently understood.

The present study addressed these issues by measuring a range of spatial and mathematics skills. Like other studies, we examined the intercorrelations of performance on these tasks, but we also used factor analysis to probe the latent structures that emerge when performance in both domains is analyzed together. Performance was measured in three age groups that span the elementary school years, thereby allowing us to examine age-related changes in these structures.

Within-Domain Structure

As in many areas of human performance, psychologists have studied the within-domain structures of spatial skill and mathematics to determine whether these are unitary or multidimensional constructs. One approach has been to use exploratory factor analysis to identify valid subdivisions within domains. However, the results of these within-domain analyses are complex and researchers have still have not reached consensus (see Mix & Cheng, 2012, for a review). Another approach has been to use theory-driven divisions to group tasks within each domain. However, although plausible, the psychological validity of these hypothesized structures have not been confirmed using factor analytic methods and are themselves the topic of much debate. Although this backdrop does not offer clear predictions regarding the structural underpinnings for the nexus of spatial thought and mathematics, it is the starting point for examining cross-domain structure, so we begin with a brief review.

The Structure of Spatial Thought

The dimensionality of spatial thought has been the subject of controversy for some time. Early factor analysts were unable to establish the existence of a spatial factor distinct from general intelligence (see Smith, 1964, for a thorough review) and even this broad division has been contentious (Carroll, 1993; Lohman, 1988; Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001). At the other end of the spectrum, it has been argued that spatial skill is not only distinct, but can be further divided into multiple subfactors (Carroll, 1993; Höffler, 2010; Linn & Petersen, 1985; Voyer, Voyer, & Bryden, 1995). However, these divisions do not appear to be clear-cut or stable, as there is a great deal of overlap and noise around the category boundaries, with divisions shifting as specific tasks are included or excluded (Carroll, 1993; Höffler, 2010; Lohman, 1988; Miyake et al., 2001). Furthermore, researchers have disagreed about the number of independent factors, what comprises each factor, and what to call them (Carroll, 1993; Höffler, 2010; Kelley, 1928; Lohman, 1979; Michael, Guilford, Fruchter, & Zimmerman, 1957; Thurstone, 1944). For example, some have argued that visual spatial working memory (VSWM) is an independent subfactor (Ackerman, Beier, & Boyle, 2005; Shah & Miyake, 1996), whereas others have questioned whether working memory is itself distinct from executive control, let alone divisible into modality-specific subabilities (i.e., verbal vs. visuospatial; Hambrick, Kane, & Engle, 2005). Still others have argued that spatial skill, VSWM, and general intelligence are so tightly interrelated that each might be viewed as a stand-in for the others, indicating that intelligence is, to a large extent, the ability to spatially manipulate mental models (e.g., Lohman, 1996). Taken together, these studies fail to yield a clear picture of the structure of spatial thought, although they suggest some intriguing possibilities.

It has been argued that one reason the existing factor analyses were inconclusive is that the structures they probed were not theory-driven (Uttal et al., 2013). There are many well-known theoretical distinctions in the spatial literature, including categorical versus coordinate representations of space (e.g., Kossly et al., 1989), near versus far perceptions of space (e.g., Cowey, Small, & Ellis, 1994), global versus local processing (e.g., Navon, 1977), and allocentric versus egocentric perspectives (e.g., Kesner, Farnsworth, & DiMattia, 1989), all of which derive support from behavioral and neurological evidence. More recently, a novel typology has been proposed based on two theoretical dimensions along which spatial tasks can differ (i.e., static–dynamic and intrinsic–extrinsic; Newcombe & Shipley, 2015; Uttal et al., 2013). These theory-driven distinctions offer an alternative approach to parsing the landscape of spatial skills. Still, it is an open question whether these distinctions mirror the latent ability structure underlying spatial performance, as one might demonstrate in a confirmatory factor analysis. That is not to say these distinctions do not reflect latent structures, but rather, that the dimensionality of spatial skill remains unclear. Indeed, the nature of these theoretical distinctions is still debated, with issues of measurement, context-specificity, and order of processing (e.g., parallel vs. sequential) yet in flux (see, e.g., Burgess, 2006; Lourenco & Longo, 2009).

All in all, with respect to the relation between spatial skill and mathematics, these studies seem to raise more questions than they answer. It is hard to know whether the cross-domain relation that has been observed repeatedly is based on general overlap involving the entire constellation of spatial skills or specific overlap involving one or more of the proposed subfactors. As we will see, the situation is further complicated by similar uncertainties in terms of mathematical thought.

The Structure of Mathematical Thought

The history of attempts to identify the substructures of mathematical thought via factor analysis parallels that for spatial skill. Investigators initially reported that mathematics skill was indistinguishable from general intelligence, and thus could not be considered a separate factor (Fouracre, 1926; Spearman & Jones, 1950; Werdelin, 1958; Wilson, 1933). However, evidence subsequently emerged to reveal a cluster of tasks that formed a uniquely mathematical factor comprised of arithmetic, procedural fluency, and numeracy (e.g., Barakat, 1951; Holzinger & Harman, 1938; Wrigley, 1958). A third factor was related to success on more conceptual math tasks, such as geometry and algebra (Barakat, 1951;

Holzinger & Swineford, 1946; Werdelin, 1958; Wrigley, 1958). This distinction between procedural and conceptual performance has since had a strong influence on research related to mathematics education and mathematics disability and continues to be a source of debate (Baroody, Feil, & Johnson, 2007; Geary, 1993; Hiebert & LeFevre, 1986; Schneider, Rittle-Johnson, & Star, 2011; Star, 2005). It is interesting to note that the conceptual factor was related to spatial visualization, offering an indication of at least one specific point of contact between domains.

Aside from the results of these factor analyses, many other theory- or task-driven divisions have been proposed for mathematics. For example, standards for mathematics instruction generally divide concepts and tasks into strands or topics, such as whole number concepts, operations/algebra, measurement, fractions, ratios, and geometry (e.g., Common Core State Standards for Mathematics [CCSS-M], National Council of Teachers of Mathematics [NCTM] Standards). These topics are further subdivided into specific tasks, such as the third-grade content standard: "Understand a fraction as a number on the number line; represent a fraction on a number line diagram" (CCSS-M, 3.NF.A.2). Just as for spatial skills, these mathematical divisions are mostly based on apparent similarities in content or operations with support from behavioral and neurological experiments, yet the dimensionality of mathematics also remains debated and the relations between these theory-driven distinctions and the latent structures underlying mathematics performance has not been tested directly, to our knowledge.

Potential Cross-Domain Relations

Our main question is whether all or only certain tasks within the domains of space and mathematics share processing, hence clarifying the latent structure that underlies their well-established correlation. The within-domain analyses offer hints as to ways this structure might take shape. For example, if the typology proposed by Uttal et al. (2013) corresponds to measurable differences in these latent structures, then it is possible only some spatial subtypes will relate to mathematics performance. Similarly, it is possible only some mathematical subtypes (e.g., numeracy, procedural knowledge, problem solving) will relate to spatial skill. However, there are many other possible outcomes. One is that specific abilities will share processing even if the latent structure of the tasks within each domain is unitary. For example, even if all the spatial tasks form a single stable factor with no apparent subdivisions, specific tasks such as mental rotation or VSWM could exhibit unique shared processing with mathematics. Another possible outcome is that both spatial skill and mathematics will form separate, unitary factors that are correlated but for which no specific tasks cross-load. In this case, one might argue that the previously established relation is based on more general shared processing, such as working memory or attention. A third possibility is that performance in the two domains overlaps so much that all measures load onto a single factor, suggesting that spatial and mathematical thought are so tightly linked they cannot be considered separate. Finally, there may be multiple shared processes, leading to a number of separate mixed domain factors, each of which is comprised of several spatial and mathematics tasks. The reason we take an exploratory approach is that the particular constellation of factor loadings is so difficult to predict. Still, we can make some reasonable guesses as to what the underlying shared processes might be, based on the demands of various mathematics and spatial tasks and the outcomes of previous research. In the following sections, we focus on three strong possibilities: (a) Spatial Visualization, (b) Form Perception, and (c) Spatial Scaling.

Spatial Visualization

Spatial visualization is the ability to imagine and mentally manipulate figures or objects in space. This skill could play a role in mathematics by helping children ground concepts or represent a problem space. For example, when children interpret a word problem, they may build a mental model to represent the problem elements and relations (e.g., Huttenlocher, Jordan, & Levine, 1994), much as they do when they are reading a story (Glenberg, Brown, & Levin, 2007). Spatially grounded mental models may also support the representation of complex mathematical relations, such as the hierarchical, nested structure of multidigit numbers (Laski et al., 2013; Thompson, Nuerk, Moeller, & Cohen Kadosh, 2013) or the part-whole relations represented in rational numbers (Matthews, Chesney, & McNeil, 2014). This view bears some relation to the notion that people ground symbolic and abstract thought in bodily movement through space (e.g., Barsalou, 2008; Lakoff, & Nunez, 2000), and thus suggests there may be particularly strong connections between spatial tasks that are based on movement and relative positions in space, such as perspective-taking, map reading, block design, and mental rotation, and mathematics tasks with relatively complex conceptualization requirements, such as interpreting word problems, comprehending place value, and fractions.

Form Perception

Form perception is the ability to recognize shapes and tell them apart, distinguish shapes from their backgrounds, and decompose them into parts. This skill could be related to mathematics in terms of its symbol reading demands. When children read mathematical symbols, they must make fine spatial discriminations, such as detecting the difference between a plus sign (+) and a minus sign (−), or noticing that 126 is different from 162 because the positions of "6" and "2" have shifted. Research has shown that adults are sensitive to these spatial relations in written mathematics and their performance can be disrupted by subtle spatial shifts (Landy & Goldstone, 2007, 2010). This shared processing might explain, at least partially, the correlations between reading and mathematics performance that have previously been attributed to semantics (Geary, 1993; Krajewski & Schneider, 2009). Although this correlation may reflect a strictly verbal component of mathematics, it might also reflect the basic visuospatial components of reading, such as form perception and contrast sensitivity (e.g., Lovegrove et al., 1982), that are also needed for reading and writing mathematical symbols. Such shared processing could be evident in mathematics tasks that require attention to subtle spatial relations in symbolic notation, such as multistep calculation, missing term problems, algebra, and interpreting charts and graphs, as well as spatial tasks that involve reproducing spatial locations and forms, such as VSWM, map reading, and figure copying. Consistent with this proposal, recent studies have found that the relations between

VSWM and mathematics shift depending on the symbol reading demands of particular tasks (e.g., addition with carrying is related to VSWM in elementary students, but addition without carrying is not; Caviola et al., 2012).

Spatial Scaling

A third possible connection could involve spatial scaling—the ability to distinguish absolute and relative distances, and recognize equivalence across different spatial scales. This ability has been linked to numeracy and symbol grounding in mathematics (Newcombe, Levine & Mix, 2016), so it is a strong candidate for cross-domain overlap with mathematics. For example, the ability to apprehend and represent spatial extent could contribute to development of the mental number line (e.g., DeHaene et al., 1993), correct placement of written numerals on a physical number line (e.g., Siegler & Opfer, 2003), and ordinal comparisons between either discrete or continuous physical quantities (e.g., Halberda & Feigenson, 2008), or between written numerals (e.g., Mix, Prather, Smith, & Stockton, 2014). If so, we might expect to find especially strong connections between mathematics tasks that focus on number meaning, such as number line estimation, and spatial tasks that require attention to relative distance or scaling, such as map reading (as some have already shown, see Barth & Paladino, 2011; Slusser, Santiago, & Barth, 2013).

Developmental Changes

Another consideration in evaluating the common structures underlying spatial skill and mathematics is whether these structures are stable across development. Although research has shown that spatial skill and mathematics are related throughout childhood and into adulthood, few studies have compared age groups on the same tasks to see whether the strength or quality of these relations changes (but see LeFevre et al., 2013, for evidence from a longitudinal approach), and none have examined task-specific cross-domain relations across age. Given that there are several distinct ways in which spatial tasks and mathematics tasks might share processing (as outlined above), it is important to examine whether the patterns of shared processing remain constant.

One reason major qualitative shifts could occur is that there are changes in mathematics content as children progress through school. For example, the shift from symbol grounding for number in the early grades to multistep equation solving in middle school might be reflected in strong relations with spatial scaling or mental rotation among younger children that are replaced by equally strong relations with visuospatial working memory and form perception among older children. Alternatively, qualitative shifts might take place depending on whether a task is novel or challenging. Some investigators have suggested that spatial skills are important for comprehending new mathematics content but become less integral once tasks are mastered or become automatic (e.g., Ackerman, 1988; Uttal & Cohen, 2012). Perhaps there is a developmental cycling wherein spatial skills are tightly linked to mathematics at each age level, but only with the specific mathematics tasks that are newly introduced at each age.

Even if the qualitative pattern of relations among spatial skills and mathematics is constant across age, there could be age-related changes in strength. Perhaps several specific spatial skills are weakly associated with mathematics at first, but these relations become more tightly connected over development (e.g., as the demands of forming complex mental models for mathematical problems and making distinctions among mathematical symbols increases). Or there could be strong relations in early childhood that gradually fade by middle school age as the demands of mathematics tasks shift and become more proceduralized. To illustrate, research using dual-task interference with adults showed that phonological, but not visuospatial interference affected performance (Logie, Gilhooly, & Wynn, 1994), whereas the mathematics performance of 6-year-olds was strongly affected by visuospatial, but not phonological disruptions (McKenzie et al., 2003). In 9-year-olds, performance was affected by disruptions in both (McKenzie et al., 2003). Taken together, these studies provide evidence for at least one such developmental shift—in this case from a strong visuospatial/mathematics link in early childhood to a less strong and finally absent link in adulthood.

Little is known about these developmental patterns because the majority of related research has focused on adolescence and adulthood. Of the relatively few studies with children, most have focused on VSWM. For example, we know that strong VSWM is related to superior performance on counting tasks (Kyttälä, Aunio, Lehto, Van Luit, & Hautamaki, 2003; LeFevre et al., 2013), number line estimation (Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Thompson et al., 2013), and nonverbal problem solving (LeFevre et al., 2010; Rasmussen & Bisanz, 2005), as well as better overall math performance (Alloway & Passolunghi, 2011; Dulaney, 2014; Gathercole & Pickering, 2000; Geary, 2013; Holmes, Adams, & Hamilton, 2008; Laski et al., 2013; Li & Geary, 2013; Meyer, Salimpoor, Wu, Geary, & Menon, 2010; Raghubar, Barnes, & Hecht, 2010). In terms of other spatial skills, there is emerging but limited research demonstrating correlations with mathematics outcomes in children. These include mental transformation (Gunderson, Ramirez, Beilock, & Levine, 2012), mental rotation (Carr, Alexeev, Horan, Barned, & Wang, 2015; Kyttälä et al., 2003; LeFevre et al., 2013), block design (Johnson, 1998; Markey, 2010), and spatial relations (Mazzocco & Myers, 2003).

Though these studies provide promising evidence for relations between mathematics and a few specific spatial skills in childhood, the nature and developmental course of these relations is not yet firmly established. Indeed, some research has failed to show these effects (e.g., Carr, Steiner, Kyser, & Biddlecomb, 2008), so the relations themselves (other than VSWM perhaps) must still be confirmed. Also, because the spatial skills in these studies were either tested in isolation or combined into single composite measures, it is hard to know whether the effects are general or specific. Furthermore, the literature on spatial development has identified a number of key abilities that have not yet been studied with respect to mathematics, such as perspective-taking, map reading, figure-ground processing, and navigation (e.g., Cornell & Hay, 1984; Huttenlocher & Presson, 1973; Liben, 2001; Marzolf & DeLoache, 1994; Newcombe & Huttenlocher, 1992; Rieser, Garing, & Young, 1994; Uttal, 2000; Shusterman, Ah Lee, & Spelke, 2008). Finally, as noted above, the existing studies tend to use a single spatial task, a single age group, and composite mathematics scores, making it difficult to discern process-level connections that may unfold over developmental time. To achieve a more comprehensive picture of how spatial skills and mathematics relate, a large-scale study is needed that includes several age groups and a broad range of subtests within each domain. The present study addresses that gap.

The Present Study

In this study, we collected multiple measures of spatial and mathematics skill in three age groups (kindergarten, third, and sixth grades). We then used these measures to evaluate the latent structures underlying performance using factor analysis and multiple regression. The measures were chosen to represent a broad sample of skills within the two domains (see Measures). Although we did not attempt to test any of the particular subrelations that have been claimed within each domain, we did include tasks that allowed us to tap into the three potential mechanisms by which spatial skill and mathematics might be related, as outlined above (i.e., spatial visualization, form perception, and spatial scaling).

Method

Participants

A total of 854 children participated. The sample was drawn from 33 schools serving a range of rural, suburban, and urban communities in the Midwestern United States (nine communities). The average free/reduced lunch rate across the nine communities was 40% (range = 0%–98%). Children's parents were contacted through their schools and only children whose parents signed an consent form approved by the institutional review board were tested. Out of 3,749 children whose parents were contacted across 33 schools, a total of 952 had parents who gave consent (i.e., 25%). Of these, 98 children were excluded because (a) tests were missing due to student absences, children who declined to participate, or schools that declined to participate after consents were turned in (n = 26), (b) tests were administered incorrectly or not recorded due to experimenter error (n = 68), or (c) children were part of a special population (English language learners, special education students, etc.; n = 4). The final sample of 854 children was divided into three age groups: kindergarten (n = 275, 131 boys, mean age = 6.04, SD = .40), third grade (n = 291, 142 boys, mean age = 9.04 years, SD = .41) and sixth grade (n = 288, 131 boys, mean age = 11.74 years, SD = .44).

Procedure

Children completed a battery of up to 15 tests (depending on age) that measured spatial skills, mathematics, and vocabulary (See Table S1 in the online supplemental material). Children were tested in three or four 1-hr sessions (depending on age) that took place over the course of two weeks. Some tests were administered in groups—either small groups (n = 4–6) for kindergarten and third-grade students, or entire classes (n = 25–30) for sixth-grade students. Other tests were administered individually. (Details provided below in the Measures section.) In the kindergarten and third-grade group tests, screens were placed between students to ensure independent work. The test order was blocked and counterbalanced by individual versus group administration. Further, the order of the tests within each block varied randomly. Children in all three grades received a decorated folder as a reward for participation.

Measures

Our rationale for choosing measures was empirical, theoretical, and developmental. We first chose a range of tasks that represented the various classes of spatial ability identified in previous factor analyses (Carroll, 1993; Höffler, 2010; Linn & Petersen, 1985; Voyer et al., 1995), as well as Uttal et al.'s (2013) more recent typology. We further ensured that specific tasks with previous evidence of a connection to mathematics were included (e.g., mental rotation, visuospatial working memory, figure copying), as well as several others for which the connection to mathematics had not yet been tested (e.g., perspective-taking, map reading).

Several of our choices were based on theoretical predictions about the potential shared processing between spatial skill and mathematics. First, we reasoned that if there is a lifelong relation between mathematics and spatial skill that predicts success in STEM careers, it must extend beyond tasks with an obvious spatial component, such as recognizing shapes or comparing lengths. So instead of targeting, for example, geometry in the early grades, we focused on core symbolic skills such as interpreting numeric symbols or manipulating symbols to perform operations and mathematical reasoning, such as solving word problems. We did, however, include number line estimation in all three grades because of its symbol grounding components (e.g., mapping multidigit numerals to ordinal meanings) and the strong relation of this task to mathematical achievement (e.g., Geary, 2011; Schneider, Grabner, & Paetsch, 2009).

Finally, there were developmental considerations. We included tasks that, insofar as possible, tapped the same conceptual skill across the grades but also were age-appropriate. For example, whole number place value is a common skill targeted by kindergarten and third-grade mathematics instruction. Most 6th-grade students have mastered whole number place value, so to measure a comparable skill, we included an age-appropriate decimal place value test (Comprehensive Mathematical Abilities Test [CMAT] Rational Numbers. We also pilot-tested all of the measures to ensure there were enough easy and challenging items at each age level to provide an even distribution of performance, and added items in some instances to manipulate difficulty (e.g., we added items with larger scale differences and rotated targets to increase difficulty in the map reading task). Finally, although we did not include geometry or measurement in the kindergarten and third-grade test batteries because this content is mostly focused on shapes and sizes (see CCSS-M), we did add geometry and graphing/data subtests in sixth grade, at which point the content involves more mathematical reasoning.

Although our goal was to survey a broad range of skills in both spatial reasoning and mathematics, we were limited in the number of skills we could include because of budgetary and practical constraints. One constraint was that, to achieve adequate statistical power, the sample size had to increase by 11 children per grade level with each additional measure. Also, more measures required longer test times per child, which was an added cost. Longer testing times were also an obstacle to recruitment and retention because of concerns about missed instructional time among parents and school personnel. Thus, it was not possible to include every measure of potential interest.

The procedures and materials for the specific measures are described below, as well as the reliabilities for each grade. Some measures were standardized tests and have published reliabilities that we report here. For the others, we computed reliabilities from our own data using Cronbach's alpha.

Mental Rotation (Adapted From Neuburger, Jansen, Heil, & Quaiser-Pohl, 2011, and Peters et al., 1995)

Two variations of Vandenberg and Kuse's (1978) mental rotation task were used. In the kindergarten/third-grade version, small groups of children were shown four figures (i.e., two-dimensional forms based on capital letters) and asked to indicate which two were the same as the target (See Figure 1). The two matching items could be rotated in the picture plane to overlap the target, whereas the two foils could not because they were mirror images of the target. In kindergarten and third grade, the task was introduced with four practice items presented on a laptop screen. Children received feedback on the correctness of their choices, and also were shown animations with the correct answers rotating to match the target. Following the practice session, children completed the 16 test items in a paper booklet (kindergarten $\alpha = .72$; third-grade $\alpha = .87$). The sixth-grade version was the same, except that children were shown 12 items consisting of perspective line drawings of three-dimensional block constructions, two of which could be rotated in the picture plane to match the target, and the practice trials were not presented on a laptop screen ($\alpha = .79$). Children received credit for answering each item correctly only if both matches were identified.

Visual Spatial Working Memory (Adapted From Kaufman & Kaufman, 1983)

On each test trial, children were shown a 14 cm × 21.5 cm grid that was divided into squares (e.g., 3 × 3, 4 × 3), with drawings of objects displayed at random positions within the grid. Item difficulty was manipulated by adding divisions to the grid (up to 5 × 5) and adding objects (up to nine). On each trial, the stimulus display was left in full view for five seconds and then it was removed and children indicated where the drawings had appeared by marking an "X" in the previously filled positions on a blank grid of the same size and shape. Note that the grids were marked with lines for response items but not the stimulus items. Stimulus displays were presented on a laptop computer and children made their responses in individual, paper test booklets. The test was introduced with two (sixth grade) or three (kindergarten and third grade) practice items for which children received feedback on the correctness of their answers and were allowed to compare their responses to the stimulus display. The test trials ($n = 19$ for kindergarten, $n = 15$ for third grade, $n = 29$ for sixth grade) began immediately after the final practice trial. The test was group administered. Because we modified the test significantly and the publication date was several decades ago, we computed reliabilities based on our own data ($\alpha = .74$, .63, and .82 for kindergarten, third grade, and sixth grade, respectively).

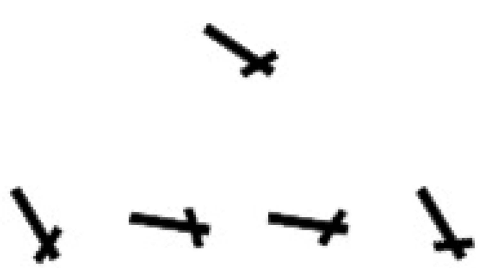

Figure 1. A sample mental rotation item (Novack, Brooks, Kennedy, Levine, & Goldin-Meadow, 2013). In this task, children were asked to circle two shapes on the bottom that match the one on the top.

Test of Visual Motor Integration (6th ed.; VMI; Beery & Beery, 2010)

On each trial, children copied a line drawing of a geometric shape on a blank sheet of paper. There were 18–24 trials, depending on the age of the child, over which the figures became increasingly complex. We administered the test in small groups. The reliability of the VMI based on a split-half correlation (reported in the test manual) was .93.

Block Design (Wechsler Intelligence Scale for Children—Fourth Edition; WISC–IV; Wechsler et al., 2004)

On each trial, children were shown a printed figure comprised of white and red sections, and they produced a matching figure using small cubes with red and white sides. The test was individually administered following the instructions in the WISC-IV manual. Items ranged in difficulty and children completed different numbers of items depending on their basal and ceiling performance. The reliability coefficient reported in the WISC-IV manual for the Block Design subtest is between .83 and .87 depending on age group.

Map Reading (Adapted From Liben & Downs, 1989)

Kindergarten and third-grade students completed 14 test trials in which they were first shown a full color three-dimensional model town with buildings, roads, a river, and trees (See Figure 2). The model was 10- × 10-in. in area and the tallest structure was 0.50 in. high. The sixth-grade task was similar except the locations were presented on full-color screenshots of three-dimensional virtual models printed on 8-1/2 × 11 sheets of paper (8 images total, one per trial). On each trial, a location was identified on the model, and children marked the same location on a two-dimensional scale map (six ×6 in.). Item difficulty was manipulated by varying the scale ratio of the map (1:1, 1:2.5) and degree of rotation between the photograph or model, and the map (0 to 180). The items were ordered from easiest to most difficult based on the results of pilot-testing. Feedback was given on the first three test questions to ensure that children understood the task. Children in the younger age groups were tested individually, but sixth-grade students completed the test in groups. The reliability of this task was kindergarten $\alpha = .56$; third-grade $\alpha = .72$; sixth-grade $\alpha = .57$.

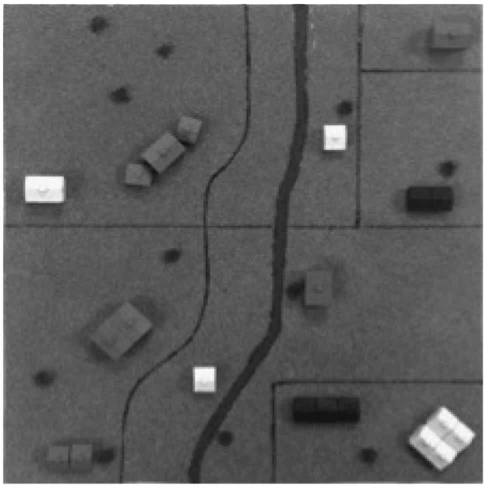

Figure 2. An overhead photograph of the three-dimensional model used in our novel map reading task. In this task kindergartners and third graders were asked to look at a position on a map and point to its corresponding position on the model.

Perspective Taking (Frick, Mohring, & Newcombe, 2014; Hegarty & Waller, 2004; Kozhevnikov & Hegarty, 2001)

The two perspective taking tasks required children to imagine a scene from different points of view. In the kindergarten/third-grade version (adapted from Frick et al., 2014), children saw a set of Playmobil figures in a particular arrangement (See Figure 3). Then, they were shown four pictures and asked to indicate which picture was taken from each character's perspective. Items varied in difficulty based on the number of objects in the pictures and the angles of view. The 27 test questions were preceded by 4 practice items with feedback. The reliability of this test was kindergarten $\alpha = .56$; third-grade $\alpha = .87$. The sixth-grade version (adapted from Kozhevnikov & Hegarty, 2001) was similar. Children were shown six to eight objects arranged in a circle. They were asked to imagine standing next to one object while directly facing another object, and then draw an arrow toward a third object to indicate their angle of view from this perspective. After two practice items with feedback, they completed the 12 test items. Responses were scored based on the number of degrees they deviated from the correct angle on each item. The reliability of this test for the sixth-grade students was $\alpha = .84$.

Place Value

We assessed place value concepts in kindergarten and third-grade students using a set of 20 items that required children to compare, order, interpret multidigit numerals (e.g., Which number is in the ones place?"), as well as match multidigit numerals to their expanded notation equivalents ($342 = 300 + 40 + 2$). Reliability on this experimenter-constructed measure was $\alpha = .79$ at kindergarten and $\alpha = .79$ at third grade. Similar concepts were assessed in sixth-grade students on the Rational Numbers subtest (CMAT; $\alpha = .94$ reported for 12-year-olds in test manual). As in the lower grades, children were asked to compare, order, and interpret written numbers, but these included a mixture of multi-digit numerals, fractions, and decimals.

Word Problems

For kindergarten and third-grade students, we tested children's problem solving ability using 12 word problems from the Test of Early Mathematics Ability–Third Edition (TEMA-3, Ginsburg & Baroody, 2003; kindergarten $\alpha = .70$; third-grade $\alpha = .63$). The TEMA-3 is a test of numerical skills, such as cardinality, calculation, and commutativity. The test was individually administered to children following the instructions in the test manual. Although children completed the entire TEMA-3 for use in another set of analyses, we analyzed only their performance on word problems here. To measure performance on word problems among sixth-grade students, we used the Problem Solving subtest from the CMAT ($\alpha = .89$ reported for 12-year-olds in test manual).

Calculation

To measure calculation, we used a group-administered test consisting of 12–28 items with age-appropriate arithmetic problems (kindergarten: $n = 16$, $\alpha = .76$; third grade: $n = 12$, $\alpha = .69$; sixth grade: $n = 28$, $\alpha = .77$). In kindergarten, the problems consisted of one- to four digit whole number addition and subtraction problems. The third-grade version also had one- to four digit whole number addition and subtraction problems, but also included four whole number multiplication and division problems (one to three digits). The sixth-grade calculation test consisted of 28 items that sampled from all four operations. Sixteen of these items used whole numbers up to five digits, and 12 of the items used decimals.

Missing Term Problems/Algebra

We analyzed children's performance on missing term problems as a separate measure because previous research suggested a causal relation between spatial skill and performance on this particular problem type (Cheng & Mix, 2014). In missing term problems, children find the solution to a calculation problem where the missing value is not the sum or difference (e.g., $X + 9 = 12$). Only kindergarten and third-grade students completed these problems because they are not challenging for most sixth-grade students ($n = 8$ items, kindergarten: $\alpha = .61$; third grade: $\alpha = .71$).

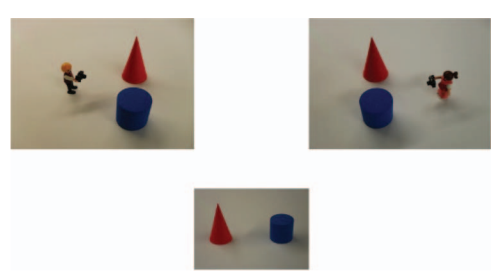

Figure 3. A sample perspective-taking item (Frick, Möhring, & Newcombe, 2014). In this task kindergarteners and third graders saw scenes of "Peter" (left) and "Lisa" (right) taking pictures of objects and then were asked to say who took a picture (bottom). See the online article for the color version of this figure.

However, we used the Algebra subtest from the CMAT to measure a similar mathematical ability in these older students (although clearly, algebra has additional components and cognitive demands). The CMAT is standardized for the age range 7 to 19 years of age and was administered in groups ($n = 10$–25). The reliability for the CMAT Algebra subtest was reported in the test manual as $\alpha = .88$ for 12-year-olds.

Number Line Estimation (Booth & Siegler, 2006; Siegler & Opfer, 2003)

Children were tested in groups of four to six. They were first shown a line with a numeral at each end (e.g., 0 and 100). Then they were shown a stimulus card with another written numeral and asked to mark where it would go on the number line. To introduce the task, the experimenter asked children to mark the approximate positions for "small" and "big" numbers, and ensured their responses were relatively correct, based on left-to-right order. They did not, however, provide feedback about the correct positions in terms of numerical magnitude. The particular numbers at the number line endpoints, and the range of stimulus values in between, varied by age group. Specifically, kindergarteners placed the numerals 4, 17, 33, 48, 57, 72, and 96 on a 0-to-100 number line (even–odd reliability: $r = .37$); third-grade students placed 3, 103, 158, 240, 297, 346, 391, and 907 on a 0-to-1000 number line (even–odd reliability: $r = .32$); and sixth-grade students placed the following on a 0-to-100,000 number line: 25,000; 61,000; 49,000; 5,000; 11,000; 2,000; 15,000; 73,000; 8,000; and 94,000 (even–odd reliability: $r = .56$). Although these reliabilities were relatively low, related work shows linear R^2 values for subsets of number line estimates vary widely (see Young & Opfer, 2011). Also, when we computed the reliabilities using error rate instead of linearity, they were well above conventionally accepted levels (kindergarten $\alpha = .74$, third grade: $\alpha = .87$, sixth grade: $\alpha = .86$). In our analyses, we focused on linearity because this variable captures internally consistent placements (i.e., sets or responses that were linear relative to each other even if they were not mapped onto the number line itself) that might be missed if absolute distance to the target were used (i.e., if all the responses were skewed to the high or low end of the number line but were nonetheless, increasing linearly); however, we also investigated whether task relations changed when error rates were used instead.

Fractions

Fraction concepts are typically introduced in third grade and become a major part of the mathematics curriculum by sixth grade (e.g., CCSS-M). For this reason, we did not include fraction items in the kindergarten test battery. In third grade, we included four items that tested fraction equivalence and simple calculation with common denominators ($\alpha = .56$). We used two measures to estimate sixth-grade students' understanding of fractions. One measure was a set of 22 items that tested comparisons, calculation with and without common denominators, and calculation with mixed numbers ($\alpha = .75$). The second measure was a version of number line estimation task in which the number line was anchored with 0 and 1, and the quantities to be placed were all fractions (i.e., 1/4, 1/19, 2/3, 7/9, 1/7, 3/8, 5/6, 4/7, 12/13, 1/2; split half reliability for linearity: $r = .40$, α for error rate $= .78$; e.g., Fazio, Bailey, Thompson, & Siegler, 2014). These two measures (fraction concepts and fraction number line) were entered separately so that we could better evaluate the relations involving number line estimation.

Supplemental Sixth-Grade Tests

The breadth of mathematics skills increases markedly in middle school and it seemed possible that skills we had not measured in younger children might be related to spatial skill in sixth grade. We therefore assessed sixth graders' performance on two subtests from the CMAT— Charts and Graphs ($\alpha = .91$) and Geometry ($\alpha = .77$) — to tap those additional skills. In the Charts and Graphs subtest, students are shown data in graphic form and asked questions that require them to interpret the information (e.g., when shown a bar graph with the number of packages mailed each day for a week, students are asked how many packages were mailed on Wednesday, and whether more packages were mailed on Friday or Monday). For Geometry, they were asked to identify geometric forms (angles, lines, solids, etc.), solve for unknown angles, calculate perimeter, area, and volume, and so forth.

Verbal Ability

To estimate and control for children's verbal skill, we used the Picture Vocabulary subtest from the Woodcock–Johnson Test of Achievement–3 (WJ-3). Although not a comprehensive assessment of intelligence, this test also provided a reasonable and easily administered estimate of general intellectual ability, based on previous studies demonstrating a strong relation between vocabulary and scores on IQ tests (e.g., Sattler, 2001; Woodcock, McGrew, & Mather, 2001). On each item, children were shown a picture and asked to name it (e.g., "What kind of insect is this?"). The test was individually administered according to the instructions in the test manual. The reported reliabilities were high (kindergarten = .73; third grade = .77; sixth grade: = .74–79).

Results

We carried out a series of analyses that probed the relational structure of these measures, both within and across domains. All analyses used children's raw scores. To control for differences in verbal ability, we specified models that used the residualized covariance matrix after partialing out children's WJ-3 Vocabulary scores.

We first calculated pairwise Pearson correlations for all the included measures. Next, we examined the factor structures both within the domains of space and math and, critically, across the two domains when all measures were entered into a common analysis. Finally, we used multiple regression analysis to determine how much variance in domain performance was accounted for by each of the cross-domain measures. Our aim was to go beyond the question of whether spatial and mathematical skills are related, to ask how they are related with respect to specific measures and processes.

Correlations Among Subtests

Before examining the overall patterns of correlation among the various measures, we equated the scales across grade levels and tasks, by transforming children's scores on each measure into z-scores. We also adjusted the threshold for significance using the Bonferroni correction for multiple comparisons. As shown in Table S2 in the online supplemental material, nearly all of the measures were significantly intercorrelated, both within and across domains. Although it is not represented in this composite table, the same basic pattern (i.e., nearly all pairwise correlations reaching significance) held within each grade as well. Thus, as shown in previous research, children who performed better on specific math and spatial tasks tended to perform better on other math and spatial tasks, suggesting underlying shared variance based on one or more overlapping abilities.

These pairwise correlations, though significant for nearly all pairs, do not necessarily mean the tasks are interchangeable. The underlying processes that connect one pair of tasks could be different from those connecting another pair, even if the relations are equally strong. Also, if there is unique variance associated with particular tasks, this would be obscured when only two tasks are compared at a time. As we will see, a more nuanced picture emerges when the various interrelations are considered simultaneously, via factor analysis and multiple regression analysis.

Next, we constructed composite scores for spatial and mathematical performance and assessed their relations to determine whether these two broad classes of ability were correlated at each grade level. To construct the composites, we averaged the z-scores across measures for each child, resulting in an individual composite score for each subbattery (i.e., spatial measures and mathematical measures). At each grade level, we tested the relation between these composites using partial correlations controlling for vocabulary skill. The correlations were significant at each grade level (kindergarten: $r = .45$; third grade: $r = .49$; sixth grade: $r = .54$, all ps $< .001$). We further examined the same scores using a confirmatory factor analysis in which performance on the spatial tasks was entered as a factor separate from performance on the mathematical tasks. Because SEM models are more sensitive to latent variables and better account for measurement error, correlations within an SEM model may be more accurate. Using this approach, with vocabulary controlled, we found even stronger correlations: kindergarten, $r = .60$; third grade: $r = .60$; sixth grade: $r = .64$, all ps $< .001$). Although the correlations at sixth grade appeared to be slightly higher than the others, none of the pairwise comparisons between grades, for either the composite scores or SEM models, were significant (Fisher's r-to-z transformation, range: $z = 0.00$ to 1.41, $p = 1.00$ to 0.16, two-tailed). Thus, there was no evidence that the overall relation between general spatial ability and general mathematics ability changes in strength with development.

Factor Analyses

Next, we report the results of three exploratory structural equation models (ESEM), two of which focused on within domain performance for spatial and mathematical tasks separately, and one cross-domain analysis in which all the subtests were considered together. All three analyses were carried out with an oblique geomin rotation in the Mplus 7.0 program (Muthén & Muthén, 1998–2012). We chose oblique rotation because it allows factors to correlate if, in fact, they do. Based on our overall correlation analysis and previous research, this outcome seemed likely. Also, because oblique rotation is nonorthogonal, it has the potential to tap both within and cross-domain loadings in the same model, even if tasks within each domain happen to be tightly correlated.

As noted above, data were submitted as raw scores for all measures, in part because scores are normalized as part of the factor analysis process so there is no need to use standard scores. As a further safeguard against non-normal distributions, we conducted all the reported factor analyses using maximum likelihood estimation with robust standard errors (i.e., MLR). MLR uses Huber sandwich estimation to provide standard errors that are robust against specification errors due to non-normal distribution (Freedman, 2006; Muthén & Muthén, 1998–2012; Wang & Wang, 2012)—an approach that has proven successful in simulation studies with distributions ranging in skewedness from −2 to 2 degrees (Chou & Bentler, 1995; Chou, Bentler, & Satorra, 1991). An examination of the distributions of scores used in the present study confirmed that all fell within this range.

For each analysis, we first extracted factors until they no longer added significant explanatory power as indicated by their eigenvalues. Specifically, we calculated 95% confidence intervals around each factor's eigenvalue based on a mathematical function that estimates error (see Larsen & Warne, 2010) and rejected models with factors for which the lower bound of the confidence interval was 1.00 or less. An eigenvalue of 1.00 indicates that a factor does not account for more variance than would an individual measure (i.e., 1 divided by the total number of variables) and a common rule of thumb is to reject models for which this is the case (i.e., the Guttman rule). However, the use of confidence intervals is more conservative and provides greater assurance in cases where eigenvalues approach 1.00. It also guarantees that the factors we retained were well above threshold in terms of explaining unique variance.

After identifying the number of informative factors, we determined the optimal rotation for each model and evaluated model fit using three indices. The root-mean-square error of approximation (RMSEA) divides estimated model error by its degrees of freedom and adjusts for sample size (Steiger, 1990). Because RMSEA estimates the "badness of fit," lower RMSEA values are better and a generally accepted cut-off is .08 (Browne & Cudeck, 1993; Hu & Bentler, 1999; Steiger, 1989). The comparative rit index (CFI) is the ratio of improvement obtained when a model generated from the data is compared to a null model that assumes no covariation among measures. A CFI greater than .95 is generally taken to indicate reasonable fit (Hu & Bentler, 1999; Raykov & Marcoulides, 2006). Standardized root mean residual (SRMR) compares an observed correlation to an ideal correlation and estimates the difference by averaging the absolute values of the correlation residuals. Like RMSEA, lower SRMR values indicate better fit and an SRMR of 0.08 or less is generally considered acceptable (Kline, 2005). Note that although such cut-offs have been debated, the risk of false rejection declines with sample sizes greater than 200 (Chen, Curran, Bollen, Kirby, & Paxton, 2008), as we achieved for each age group in the current study.

Once the model with the best fit was identified, we determined which tasks loaded onto each factor significantly following previously established procedures (Cudeck & O'Dell, 1994; Schmitt & Sass, 2011). Specifically, we derived z-values by dividing the factor loading for each measure by its standard error. Only tasks with z-values greater than 1.96 were considered significant ($p = .05$).

Spatial Factor Analysis

For the spatial tasks alone, we found convergence on a single factor at each grade level. That is, the first factor in each grade level had adequate eigenvalues (kindergarten = 2.35 (1.96; 2.74); third grade = 2.70 (2.26; 3.14); sixth grade = 3.19 (2.67; 3.71), but the second factors did not (kindergarten = 0.97 (0.81; 1.13; third grade = 0.85 (0.71; 0.99); sixth grade = 0.75 (0.63; 0.87). Furthermore, the fit of these one-factor models was good by all three indices (kindergarten: RMSEA = 0.01 (0.00; 0.07), CFI = 1.00. SRMR = 0.02); third grade: RMSEA = 0.03 (0.00; 0.07), CFI = 1.00, SRMR = 0.02); sixth grade: RMSEA = 0.04 (0.00; 0.08), CFI = .99, SRMR = 0.02).

As shown in Table 1, all the spatial measures loaded significantly onto this factor at each grade level, suggesting that these spatial skills were broadly overlapping and perhaps best considered a unitary construct. This outcome is unexpected given previous claims of subfactors among spatial tasks (Carroll, 1993; Höffler, 2010; Voyer et al., 1995) as well as the various theory-driven subdivision that have guided research on spatial ability (e.g., Uttal et al., 2013). One explanation might be that many of the previous studies reporting multiple spatial factors have used older age groups (adolescents and adults; e.g., Voyer et al., 1995). If these distinctions are weak in younger children, then they may have been missed in our elementary aged sample. Also, as we noted earlier, we were not able to include as many spatial measures as one might need to fully evaluate the internal factor structure within each domain. Still this outcome was surprising given that the set of spatial measures we included was rather diverse.

The finding of a single spatial factor at all three ages might seem to suggest there is enough shared variance that any spatial task would have as strong a relation to mathematics as any other. However, although well fitted to the data, this factor did not account for 100% of the variance of any measure. The task with the highest loading was Block Design in sixth grade, and, even in this case (with a loading of .74), the common factor accounted for only 55% of the variance in performance. Thus, even though all of these spatial tasks may tap the same latent ability, they must tap other capacities as well—capacities that do not cohere into additional separate factors but which may have shared variance with mathematics nonetheless.

Mathematical Factor Analysis

When only the mathematics measures were considered, we again found that a one-factor model was the best fit in all three grades. Specifically, the first factor in each grade level had adequate eigenvalues (kindergarten = 2.76 [2.30; 3.22]; third grade = 2.95 [2.47; 3.43]; sixth grade = 4.53 [3.79; 5.27]), but the second factor did not (kindergarten = 0.74 [0.62; .86]; third grade = 0.83 [0.70; 0.97]; sixth grade = 0.97 [0.81; 1.13]. The fit of these one-factor models was generally good, with the exception of the RMSEA for kindergarten, which was slightly higher than our preferred cut-off of .08 (kindergarten: RMSEA = 0.10 [0.05; 0.15], CFI = 0.98. SRMR = 0.02; third grade: RMSEA = 0.07 [0.03; 0.10], CFI = 0.98. SRMR = 0.03; sixth grade: RMSEA = 0.05 [0.03; 0.08], CFI = 0.99. SRMR = 0.02). As shown in Table 2, all of the mathematics tasks loaded significantly onto the single factor at all three grade levels. However, as for the spatial tasks, this factor did not account for 100% of the variance in mathematics performance as none of the tasks loaded completely onto it (i.e., the highest loading tasks—calculation in kindergarten and word problems in sixth grade—accounted for only 56% of the variance).

Table 1
Within Domain Variable Loadings by Grade: Spatial Skill

Predictor	Kindergarten	3rd grade	6th grade
Mental rotation	.434* (.054)	.546* (.049)	.615* (.042)
VSWM	.449* (.056)	.531* (.050)	.606* (.042)
VMI	.687* (.045)	.384* (.064)	.567* (.044)
Block design	.571* (.055)	.722* (.045)	.739* (.036)
Map reading	.562* (.048)	.537* (.049)	.556* (.050)
Perspective taking	.245* (.068)	.594* (.047)	.609* (.042)

Note. VSWM = visual spatial working memory; VMI = Test of Visual Motor Integration.
* $p < .05$.

Table 2
Within-Domain Variable Loadings by Grade: Mathematics

Predictor	Kindergarten	3rd grade	6th grade
Place value	.604* (.046)	.615* (.047)	.687* (.032)
Word problems	.712* (.042)	.657* (.051)	.747* (.032)
Calculation	.747* (.037)	.650* (.040)	.621* (.041)
Missing terms	.531* (.057)	.702* (.038)	.623* (.036)
Whole number line (R^2)	.486* (.044)	.406* (.059)	.257* (.068)
Fractions		.512* (.048)	.651* (.036)
CMAT-Charts			.693* (.033)
CMAT-Geometry			.551* (.040)
Fraction number line (R^2)			.455* (.048)

Note. CMAT = Comprehensive Mathematical Abilities Test.
* $p < .05$.

Cross-Domain Factor Analysis

Our central question was whether, when entered into a common model, spatial and mathematics measures would converge into common factors based on shared variance across domains. Given the strong internal factor structures within each domain, one likely outcome would be to replicate the within domain factor models and observe no cross-domain loading. However, as we will see, there was enough shared variance across domains to result in novel factor structures when the two domains were combined.

At all three grade levels, the first two factors had adequate eigenvalues (kindergarten: Factor 1 = 3.76 [3.13; 4.39], Factor 2 = 1.43 [1.19; 1.67]; third grade = Factor 1 = 4.21 [3.53; 4.90], Factor 2 = 1.47, [1.23; 1.71]; sixth grade = Factor 1 = 6.09 [5.08; 7.06], Factor 2 = 1.72 [1.44; 2.00]), whereas the third factor did not (kindergarten = 1.00 [.83; 1.17]; third grade = 0.89 [0.75; 1.04]; sixth grade = 0.97 [0.81; 1.13]). The fit of the two-factor models was good at each grade level (kindergarten: RMSEA = 0.05 [0.03; 0.08], CFI = 0.97. SRMR = 0.03; third grade: RMSEA = 0.05 [.02; 0.06], CFI = 0.98. SRMR = 0.03; sixth grade: RMSEA = 0.04 [0.03; 0.06], CFI = 0.98. SRMR = 0.02).

At each grade level, one of these factors was primarily spatial and the other was primarily mathematical (see Table 3); however, the factors were not mutually exclusive and there were several significant cross-domain loadings. In kindergarten, performance on both mental rotation and Block Design loaded significantly

Table 3
Cross-Domain Variable Loadings by Grade (Oblique Rotation)

Variable	Kindergarten Factor 1	Kindergarten Factor 2	3rd grade Factor 1	3rd grade Factor 2	6th grade Factor 1	6th grade Factor 2
Mental rotation	.296* (.074)	.248* (.075)	.500* (.076)	.087 (.089)	.616* (.050)	−.001 (.032)
VSWM	.447* (.075)	.027 (.085)	.496* (.075)	.074 (.082)	.528* (.059)	.143* (.062)
VMI	.691* (.088)	−.007 (.064)	.349* (.084)	.087 (.079)	.497* (.060)	.125* (.065)
Block design	.475* (.076)	.152* (.073)	.729* (.083)	−.034 (.093)	.748* (.050)	−.009 (.049)
Map reading	.577* (.081)	−.003 (.053)	.544* (.058)	−.006 (.040)	.563* (.065)	−.011 (.059)
Perspective taking	.162* (.079)	.133 (.093)	.595* (.064)	.009 (.054)	.566* (.059)	.071 (.066)
Place value/rational numbers (CMAT)	.038 (.079)	.588* (.067)	.152 (.099)	.532* (.079)	.148* (.056)	.605* (.044)
Word problems	.002 (.039)	.700* (.052)	.135 (.105)	.576* (.087)	.065 (.064)	.709* (.053)
Calculation	−.006 (.058)	.759* (.052)	−.008 (.012)	.662* (.040)	−.019 (.059)	.637* (.054)
Missing terms/algebra	−.126 (.107)	.602* (.088)	−.043 (.087)	.729* (.065)	.134* (.064)	.545* (.053)
Whole number line (R^2)	.024 (.091)	.466* (.068)	.085 (.094)	.358* (.089)	−.045 (.075)	.285* (.070)
Fractions			.083 (.084)	.475* (.076)	−.008 (.045)	.658* (.045)
CMAT-Charts					−.018 (.060)	.702* (.053)
CMAT-Geometry					.083 (.064)	.506* (.050)
Fraction number line (R^2)					.005 (.061)	.456* (.060)

Note. VSWM = visual spatial working memory; VMI = Test of Visual Motor Integration; CMAT = Comprehensive Mathematical Abilities Test.
* $p \le .05$.

onto the mathematics factor. In sixth grade, there also were spatial tasks that loaded significantly onto the mathematics factor, but they were VSWM and VMI. In addition, in sixth grade, two mathematics tests—algebra and place value (as measured on the Rational Numbers subtest of the CMAT)—cross-loaded significantly onto the spatial factor.[1] In third grade, there were no significant cross-loadings; however, as we will see in the next set of analyses, this may have been due to a broad distribution of variance across spatial tasks as they relate to mathematics in this age group. Overall, our results indicate that rather than simply replicating the factor structure that emerged in each domain, there were significant patterns of cross-domain loading and these patterns differed across grade levels.

One could argue that these cross-loadings provide only weak evidence for a connection between spatial skill and mathematics because these effects are small and limited to a few tasks. However, the cross-domain loadings do not represent the entire shared variance between spatial skill and mathematics. Recall that the model allowed the factors to correlate and, as expected, the correlations were high (kindergarten: $r = .50$; third grade: $r = .50$; sixth grade: $r = .53$). In the context of the overall factor structure, the cross-domain loadings indicate unique or particularly strong overlap for particular tasks, above and beyond the variance that is shared generally across the two factors.

But what general shared processing might these strong interfactor correlations reflect? To explore the nature of these relations, we repeated the exploratory factor analysis using orthogonal rather than oblique rotation. Whereas oblique rotation can highlight cross-loadings by taking more general interfactor correlations into account, orthogonal rotation prevents factors from correlating, so all shared variance is expressed in the factor structure itself. As shown in Table 4, the first factor in the orthogonal analysis was comprised of significant loadings on all of the tasks from both domains—essentially a unitary, general factor. This pattern was obtained in all three grades. The second factor was mostly spatial. Indeed, it was exclusively spatial in kindergarten and third grade, and all of the spatial tasks loaded onto it significantly. In sixth grade, this second factor also was comprised of all the spatial tasks, but included two mathematics tasks as well—place value and algebra—just as we found in the cross-loadings for the oblique rotation model. This finding reinforces the notion that, of the mathematics skills we tested, place value and algebra are particularly sensitive to differences in spatial skill.

Taken together, these results indicate the majority of shared variance between spatial skill and mathematics is very general and not reducible to a particular skill or set of skills. This shared variance may reflect individual differences in general ability. Although we controlled for general ability by partialing out children's vocabulary scores, one could argue this control addressed only crystallized abilities, and might have allowed fluid abilities to covary with spatial skill and mathematics. Another possibility is that mathematical thought is inherently spatial, in the same sense that others have argued much of abstract thought is inherently spatial (Colom, Contreras, Botella, & Santacreu, 2002; Lohman, 1996), and it is for this reason that relations among the three are so difficult to disentangle. In light of this, the cross-domain loadings take on new meaning. They indicate there is task-specific shared variance over and above the strong general relation to which these same tasks contribute. Although all the spatial and mathematics

[1] One concern regarding our comparisons across grades could be that not every subskill was measured in all three grades. Specifically, three measures used in sixth grade were not tested in kindergarten or third grade: (a) Charts and Graphs (CMAT); (b) Geometry (CMAT); and (c) Fraction Number Line. Perhaps including these tasks shifted the other factor loadings in a way that would be confusable with bona fide developmental changes. When we repeated the sixth-grade cross-domain factor analysis with these tasks removed, the same factor structure and model fit were obtained on the remaining measures. The factor loadings for spatial tasks onto the mathematics factor were also the same, but the cross-loadings for place value and algebra onto the spatial factor were no longer significant ($p = .55$ and $.72$, respectively). However, when we repeated the multiple regression analyses with these tasks removed, place value and algebra were still highly correlated with the spatial factor (Place Value: $\beta = .28$, $sr^2 = .03$, $p < .0001$; Algebra: $\beta = .17$, $sr^2 = .01$, $p = .007$).

Table 4
Cross-Domain Variable Loadings by Grade (Orthogonal Rotation)

Variable	Kindergarten Factor 1	Kindergarten Factor 2	3rd grade Factor 1	3rd grade Factor 2	6th grade Factor 1	6th grade Factor 2
Mental rotation	.394* (.060)	.258* (.065)	.329* (.070)	.440* (.068)	.325* (.058)	.523* (.053)
VSWM	.248* (.068)	.389* (.068)	.313* (.071)	.436* (.067)	.421* (.050)	.449* (.053)
VMI	.334* (.061)	.601* (.081)	.256* (.073)	.308* (.079)	.387* (.057)	.423* (.053)
Block design	.387* (.064)	.414* (.069)	.318* (.087)	.638* (.071)	.385* (.058)	.635* (.050)
Map reading	.281* (.061)	.502* (.073)	.256* (.079)	.476* (.066)	.286* (.060)	.478* (.063)
Perspective taking	.213* (.081)	.141* (.069)	.296* (.084)	.521* (.069)	.369* (.055)	.481* (.052)
Place value/rational numbers (CMAT)	.607* (.046)	.035 (.070)	.606* (.055)	.142 (.096)	.683* (.033)	.128* (.052)
Word problems	.701* (.043)	.004 (.044)	.641* (.060)	.128 (.101)	.744* (.034)	.058 (.057)
Calculation	.756* (.037)	−.003 (.039)	.658* (.041)	.004 (.016)	.627* (.042)	−.014 (.050)
Missing terms/algebra	.540* (.059)	−.108 (.089)	.708* (.045)	−.025 (.079)	.616* (.038)	.115* (.056)
Whole number line (R^2)	.477* (.045)	.022 (.079)	.399* (.063)	.081 (.085)	.262* (.068)	−.037 (.064)
Fractions			.515* (.052)	.081 (.077)	.654* (.037)	−.005 (.034)
CMAT-Charts					.693* (.034)	−.013 (.052)
CMAT-Geometry					.549* (.040)	.072 (.055)
Fraction number line (R^2)					.454* (.049)	−.002 (.055)

Note. VSWM = visual spatial working memory; VMI = Test of Visual Motor Integration; CMAT = Comprehensive Mathematical Abilities Test.
* $p \leq .05$.

tasks we measured are interrelated, the cross-loading tasks stand out because they contribute extra shared variance, perhaps via multiple routes, and because these tasks contribute in both ways (generally and specifically), they may constitute the strongest contact points between spatial skill and mathematics.

It was surprising that number line estimation did not load significantly onto the spatial factor, given previous research (e.g., Gunderson et al., 2012; Thompson et al., 2013) and one might wonder whether this had to do with our choice to use linearity as the dependent measure rather than error rate. In past research, the two measures have been used interchangeably (Opfer & Siegler, 2007; Siegler & Booth, 2004); however, as we noted earlier, the linearity measure is more relaxed in the sense that it requires only correct ordinal positioning and equal spacing within the probed numbers but does not require an accurate mapping to the scale of the number line stimulus. To find out whether this difference might alter the factor loadings, we repeated the grade specific cross-domain factor analyses using error rate for number line estimation. The factor loadings for number line estimation were unchanged by this manipulation—they remained significant for the mathematics factor and not for the spatial factor.[2]

Cross-Domain Multiple Regression Analyses

The factor analyses identified latent structures that underlie performance on the spatial and mathematics tasks we measured, but they do not address directional hypotheses involving the two domains. Although we cannot establish causality in the present study, it is possible to evaluate specific predictive relations using multiple regressions. Such tests, being more sensitive and targeted than those used in exploratory factor analysis, could contribute important information needed to understand these relations. Also, compared to pairwise correlations, multiple regressions are advantageous because they take into account the intercorrelations among independent variables while also indicating the relative strength of each predictor.

We used three regression models to evaluate the cross-domain relations among spatial ability and mathematics. The first two models used the individual subtest scores from one domain to predict the factor scores generated for the other domain, and vice versa. In the third, we asked whether the relation to spatial tasks differed for new versus familiar mathematics content. In all three analyses, we controlled for verbal skill by including children's WJ-3 Vocabulary scores as one of the independent variables. As before, raw scores were used in all analyses. As a safeguard against non-normal distribution of scores, we examined the error variance for each of the reported models and confirmed that they were random and normally distributed (Raykov & Marcoulides, 2008).

Spatial Measures Regressed on Mathematics

The results of this analysis, presented in Table 5, replicated and extended the findings of the cross-domain factor analysis. First, the spatial measures as a group were significant predictors of mathematical factor scores in each grade—kindergarten: $F(7, 267) = 25.58$, $R^2 = .40$; third grade: $F(7, 283) = 23.57$, $R^2 = .37$; sixth grade: $F(7, 280) = 48.57$, $R^2 = .55$. Second, the same individual measures that cross-loaded significantly in the cross-domain factor analysis also were significant predictors in the regression analysis (See Table 5). In kindergarten, mental rotation once again had the strongest relation with mathematics, accounting for roughly 4% of

[2] With respect to the overall models, the cross-domain loading patterns were mostly the same whether error rate or linearity was used as the dependent measure, with two exceptions. First, in kindergarten, the previously significant cross-domain loading of Block Design onto mathematics became marginally significant ($p = .08$) when error rate was used. Second, in sixth grade, the previously significant cross-domain loading of algebra onto the spatial factor remained relatively high but was no longer significant when absolute error rather than linearity was used ($p = .11$). These differences appear due to slightly higher standard error for the factor loadings in the error rate model.

Table 5
Regression of Spatial Variables Onto the Mathematics Factor, by Grade

Variable	Kindergarten β	Kindergarten sr^2	3rd grade β	3rd grade sr^2	6th grade β	6th grade sr^2
Mental rotation	.235*	.041*	.154*	.016*	.049	.001
Visuospatial working memory	.084	.005	.144*	.015*	.201*	.023*
Visuomotor integration	.064	.003	.120*	.012*	.158*	.015*
Block design	.188*	.023*	.090	.004	.093	.003
Map reading	.062	.003	.067	.003	.054	.002
Perspective taking	.087	.007	.126*	.010*	.117*	.007*
WJ-3 Vocabulary	.242*	.047*	.213*	.037*	.333*	.086*

Note. Standardized coefficients (β) and unique variance accounted (squared semipartial correlation: sr^2) for when regressing math factor scores on spatial variables. WJ-3 = Woodcock–Johnson Test of Achievement–3.
* $p < .05$.

the variance. Block Design also was a significant predictor at this age, as before, explaining roughly 2% of the variance. Recall that in third grade, the factor analysis failed to reveal any significant cross-domain loadings. However, the regression analysis revealed several significant effects involving mental rotation, VSWM, VMI, and perspective-taking. It is interesting to note that these effects were smaller and more distributed across tasks than those in either kindergarten or sixth grade, and this might explain why these relations were not evident in the factor analysis. Another notable point of comparison was that mental rotation was significantly correlated with mathematics in both kindergarten and third grade, but not in sixth grade. For sixth-grade students, as indicated by the factor analysis as well, the spatial tasks most strongly associated with mathematics were VSWM and VMI. However, the multiple regression indicated a small, but significant effect of perspective-taking as well, echoing the same effect in third grade. These findings suggest a developmental transition in the relations among spatial tasks and mathematics, starting with one set of relations and moving to a different set of relations by way of a transition period that implicates them both (i.e., the third-grade predictors included all of the kindergarten and sixth-grade predictors, except for Block Design).

Mathematics Measures Regressed on Spatial Ability

The results for the second model, in which mathematical variables were regressed onto the spatial factor scores, are shown in Table 6. Mathematical variables were significant predictors of the space factor score in each grade—kindergarten: $F(6, 268) = 23.43$, $R^2 = .34$; third grade: $F(7, 283) = 24.35$, $R^2 = .38$; sixth grade: $F(10, 277) = 31.16$, $R^2 = .53$. In particular, two mathematics skills were consistent predictors of spatial skills across age: (a) interpretation of number meanings, as measured by either our place value test (K, 3rd) or the CMAT Rational Numbers subtest (6th) and (b) the ability to solve word problems.

The other predictors of spatial ability were grade-specific. In kindergarten, calculation had the strongest relation with spatial ability. This is consistent with Cheng and Mix's (2014) finding that training on a 2-D mental rotation task had a significant effect on first grade calculation scores. It is noteworthy, however, that missing term problems did not show particularly strong relations with spatial ability, as they had previously (Cheng & Mix, 2014). One reason might have been generally worse performance on missing term problems in the younger (i.e., kindergarten) children tested here ($M = 17\%$ correct) than the first grade students trained by Cheng and Mix (pretest $M = 39\%$ correct).

In third grade, performance on the fractions subtest was a predictor of spatial abilities and in sixth grade, it was algebra. Perhaps these grade-specific effects are due to the novelty of these concepts at each age point. Indeed, multidigit calculation, fractions, and algebra are well-known obstacles in mathematical development that first emerge in kindergarten, third grade, and sixth grade, respectively, so it is interesting that they are particularly sensitive to spatial ability, consistent with the novel-familiar hypothesis advanced previously (Ackerman, 1988; Uttal & Cohen, 2012).

In summary, these regressions of individual measures from each domain onto each of the domain-specific factors demonstrate that even though some correlations were not large enough to suggest a

Table 6
Regression of Mathematics Variables Onto the Spatial Factor, by Grade

Variable	Kindergarten β	Kindergarten sr²	3rd grade β	3rd grade sr²	6th grade β	6th grade sr²
Place value/rational numbers (CMAT)	.175*	.014*	.211*	.025*	.285*	.030*
Word problems	.149*	.010*	.184*	.018*	.159*	.008*
Calculation	.214*	.021*	.050	.002	.019	0
Missing terms/algebra	−.026	0	−.002	0	.158*	.011*
Whole number line (R^2)	.063	.003	.067	.003	−.011	0
Fractions			.153*	.016*	−.012	0
CMAT-Charts					−.032	0
CMAT-Geometry					.116	.006
Fraction number line (R^2)					−.032	.001
WJ-3 Vocabulary	.168*	.021*	.199*	.032*	.237*	.037*

Note. Standardized coefficients (β) and unique variance accounted (squared semipartial correlation: sr^2) for when regressing spatial factor scores on math variables. CMAT = Comprehensive Mathematical Abilities Test; WJ-3 = Woodcock–Johnson Test of Achievement–3.
* $p < .05$.

single underlying factor, there may be characteristics of specific spatial tasks that make them especially useful predictors. It is clear that some spatial tasks predicted unique variance within the general factor that underlies mathematical ability, and vice versa. Thus, although the patterns we obtained are generally consistent with those observed in the more conservative factor analyses, they go further to uncover additional measures that might have particularly strong cross-domain ties.

Novel Versus Familiar Content

As noted above, it has been hypothesized that spatial processing may be implicated more during the initial stages of learning a particular mathematical topic than at later stages (Ackerman, 1988; Uttal & Cohen, 2012). For example, when children are first introduced to fractions, they may recruit spatial processes to (a) map the symbolic representations for fractions onto spatial referents or (b) mentally manipulate fraction symbols so as to align them with more familiar whole number meanings but these relations may decrease as skills become more automatic or procedural. If so, then we might find stronger relations among spatial skills and mathematics for new versus familiar content.

To find out, we conducted a third regression analysis in which we first divided the mathematics items within each measure into two categories—new and familiar—using the grade level content standards in the CCSS-M (National Governors Association Center for Best Practices and Council of Chief State School Officers, 2010). Items related to content from the children's current grade level or higher were considered new, and items related to content from any of the preceding grade levels were considered familiar. From these items, we derived factor scores that were then regressed against each of the spatial tasks. Although the CCSS-M have not been fully implemented, they are roughly consistent with previous grade level guidelines (e.g., NCTM) and provide a reasonable approximation of the content to which children in our study had likely been exposed.

As shown in Table 7, although some of the same spatial predictors emerged as in the overall analyses reported above, there were distinct patterns for new and familiar mathematics content. In kindergarten, Block Design and perspective-taking were related only to new content. Although mental rotation was related to both new and familiar mathematics content, the relation was stronger for new content. Also, VMI emerged as a predictor in kindergarten, but only for familiar content. In kindergarten, one of the main shifts from familiar to novel content was from single to multidigit numerals (instantiated in ordinal judgments, calculation, word problems, etc.). Thus, it may be the case that spatial visualization is more important for interpreting multidigit quantities, perhaps because place value relations must be evaluated (consistent with recent research by Laski et al., 2013, discussed above). In contrast, attention to form (VMI) may be more important to familiar content as processing numerals becomes more automatic for single-digit problems in kindergarten.

In third grade, mental rotation and VSWM predicted performance on both new and familiar mathematics content, but VMI and perspective-taking were predictive for familiar content only. There also was a marginally significant relation between Block Design and novel content in this grade. Novel content in third grade consisted of multiplication, division, and fractions, whereas familiar content consisted of addition, subtraction, and multidigit numeral interpretation. As in kindergarten, it appears that attention to shapes and figure copying (as measured on the VMI) may become more important once skills are automatic. However, the other results were more mixed. Perhaps there is greater variability among children this age in terms of mastery of content, such that our definitions of novel and familiar were not entirely accurate for some students. There may also be a combination of developmental and content related processes driving these age-related changes. That is, some third-grade students may be more likely to spatially ground relations in multidigit calculation, regardless of the novelty of particular operations, whereas others, those who are more fluent with these calculations, may no longer do so.

For sixth-grade students, new content included geometric equations, translating among fractions/decimals/percentages, algebra, and word problems based on ratios, area, and proportions. Familiar content included number line estimation, calculation, interpreting charts and graphs, and fraction calculation. For the most part, there was not a differentiation between these content areas in terms of correlations with spatial tasks. That is, VSWM and the VMI were predictive of both familiar and new content, perhaps because these spatial skills support symbol reading and manipulation, which are central to all mathematics content in sixth grade, much as we saw in third grade. It is interesting that a small but significant relation between mathematics and Block Design emerged for the first time in this grade level, and was only significant for novel content. This

Table 7
Regression of Spatial Variables on Familiar and Novel Mathematics Skills

Predictor	Kindergarten		3rd grade		6th grade	
	Familiar	Novel	Familiar	Novel	Familiar	Novel
Mental rotation	.121*	.230*	.157*	.131*	.024	.068
VSWM	.083	.078	.172*	.143*	.196*	.235*
Visuomotor integration	.184*	.004	.123*	.061	.159*	.141*
Block design	.098	.180*	.018	.147*	.087	.130*
Map reading	.048	.063	.064	.029	.075	.006
Perspective taking	.040	.140*	.165*	.066	.128*	.087
WJ-3 vocabulary	.227*	.228*	.191*	.147*	.261*	.296*

Note. Standardized coefficients (β). VSWM = visual spatial working memory; VMI = Test of Visual Motor Integration; WJ-3 = Woodcock–Johnson Test of Achievement-3.
* $p \leq .05$.

provides a hint that spatial visualization may still come into play when children are learning representationally complex new topics, such as algebra, even if it is not generally associated with achievement at this age. There also was a weak relation with perspective-taking, for familiar content only.

In summary, the specific relations between spatial tasks and mathematics differed somewhat depending on whether the mathematics content was novel or familiar. Though these patterns were not completely regular across age, and some tasks appear to be implicated in both novel and familiar tasks, there were a few notable consistencies. First, Block Design was only associated with novel content (in both kindergarten and sixth grade). Second, VMI was only associated with familiar content (in all three grades). This contrast echoes the developmental shift we observed in the cross-domain factor analysis, from strong relations between spatial visualization and mathematics in kindergarten toward strong relations between form perception and mathematics in sixth grade, suggesting that this developmental relation may be recapitulated within age groups depending upon the novelty of the mathematics content. Perhaps children recruit spatial visualization when tasks are new and require grounding or conceptualization, but rely more on form perception when for rapidly reading symbols and equations when tasks are more automatic and procedural. (See Stieff, 2013, for such a pattern in chemistry students.) If so, then the overall developmental picture may be quite complicated, because this cycling pattern for novel and familiar content may occur against a backdrop of more general age-related changes in mathematics content, procedural skill, and spatial skill. This might explain why the present analysis was not entirely clear-cut and points to the need for more research to sort out these potentially complex interactions.

Discussion

The present study investigated the well-established relation between spatial ability and mathematics by examining interconnections among specific tasks both within and across domains. This approach allowed us to determine whether the relation that has been demonstrated previously is due to general overlap in processing between these domains or specific overlap in processing between particular space and mathematics tasks. We further assessed developmental changes in these specific relations by comparing patterns cross-sectionally, from kindergarten to sixth grade, and by comparing relations involving new and familiar content.

Our findings indicate strong within-domain factor structures for both space and mathematics based on the findings that (a) a one-factor solution best described the latent structures within each domain and (b) the two-factor model that best described the cross-domain structure had factors comprised of all the spatial and mathematics tasks, respectively. Contrary to previous studies reporting multidimensional structures for spatial skill and mathematics separately, we obtained no such evidence. Notably, there was no evidence that spatial skill was subdivided in terms of constructs such as spatial visualization or along the lines of the theoretical distinctions such as static versus dynamic. Similarly, we found no evidence for a procedural and conceptual factors in mathematics or task-specific groupings (e.g., whole number operations vs. fractions). That said, we did not construct our analyses or our measures in a way to probe directly for such dimensionality. Perhaps if we had included more measures or item level analyses, such distinctions would have emerged. As noted above, the number of measures we could include was limited by practical considerations and, in order to cover both domains adequately, we could not probe either domain very deeply. Still, it was surprising that no within-domain differences emerged.

We also found a strong, consistent correlation between domains across the three grade levels. In fact, when we used an orthogonal model that prevented the factors from correlating, spatial skill and mathematics formed a single, shared factor. Thus, there is a great deal of overlapping variance in the two domains, irrespective of task-specificity, that could be due to either general ability (e.g., fluid intelligence) or very basic shared processing (i.e., mathematics itself being fundamentally spatial), as some theorists have argued (e.g., Lohman, 1996). This general overlap may go a long way toward explaining the previously reported effects at the behavioral and neural levels of analysis involving spatial skill and mathematics (e.g., Dehaene et al., 1993; Hubbard et al., 2005; McKenzie et al., 2003; Walsh, 2003).

Yet, against this backdrop of highly correlated domains, there also were significant cross-domain relations involving particular tasks that varied across grade. In kindergarten, the spatial skills of mental rotation and Block Design were strongly related to mathematics. In sixth grade, the spatial skills of VSWM and form perception (VMI) were strongly related to mathematics and the mathematics tasks of algebra and place value were strongly related to spatial processing. In third grade, there were no significant cross-domain loadings in the factor analysis, but there were a number of weak but significant correlations revealed in the multiple regression analyses. It is interesting to note that these correlations comprised all of the spatial skills (except Block Design) that were significantly cross-loaded in kindergarten and sixth grade, suggesting that in terms of spatial skill and mathematics, third grade may represent the middle of a qualitative shift.

In the following sections, we return to several key issues raised in the introduction in light of our results. Specifically, we revisit the three candidates for shared processing (Spatial Visualization, Form Perception, and Spatial Scaling). We also consider possible explanations for the age-related shifts we observed.

Spatial Visualization

Of the three potential types of processing overlap we identified in the introduction, our data were most indicative of a connection between mathematics and spatial visualization. This may reflect shared processing that supports spatially grounded representations of complex relations and problem space, consistent with the embodied view of cognition (e.g., Barsalou, 2008; Lakoff & Nunez, 2000). One indication was that the mathematics tasks with the strongest relations to spatial ability—place value (all grades), word problems (all grades), fractions (third grade, sixth grade), and algebra (sixth grade)—are well-known obstacles in elementary mathematics thought to pose particularly complex representational challenges. For example, it takes years for children to unpack the syntactic structure of place value (e.g., $429 = 4x100 + 2x10 + 9x1$) and carry out multidigit procedures based on accurate interpretations of base-10 structure (Cauley, 1988; Cobb & Wheatley, 1988; Fuson & Briars, 1990; Jesson, 1983; Kamii, 1986; Kouba et al., 1988; Labinowicz, 1985; Moeller, Pixner, Zuber, Kaufmann, & Nuerk, 2011; Resnick & Omanson, 1987). Spatial skill may be related to place value understanding in particular because it helps children develop a basic level representation of these complex, problems (e.g., Laski et al., 2013). Similarly, when children interpret a word problem, they may build a mental model to represent the problem elements and relations, much as they do when they are reading a story (Glenberg et al., 2007). Spatially grounded mental models might help them conceptualize the part-whole relations in rational numbers (Stafylidou & Vosniadou, 2004) or support the relational view of equivalence that underlies balancing of algebra equations (Knuth, Stephens, McNeil, & Alibali, 2006). Another indication of this shared process is that most of the spatial tasks with significant cross-domain relations— mental rotation (kindergarten, third grade), perspective-taking (third and sixth grades) and WISC-IV Block Design (kindergarten)—have strong spatial visualization components. They require imagining movements in space, shifting and maintaining multiple orientations, or analyzing parts and wholes—the kinds of spatial skill that seem likely to help one construct a mental model or keep track of related parts.

Form Perception

A second type of shared processing we identified involved accurate perception of form, spatial layout and the like, and its potential relations with reading symbols and keeping track of steps in a complex procedure. Our results provided evidence for this shared processing as well, particularly in the later elementary grades. One indication was that spatial tasks requiring accurate memory of location (VSWM) and detailed forms (VMI) had relatively strong relations to mathematics in third and sixth grades. These tasks emerged as unique predictors in the third-grade multiple regressions and then exhibited stronger relations in sixth grade with significant cross-loadings in the cross-domain factor analysis. In light of this, it is also telling that algebra was the strongest mathematics cross-loading in sixth grade, as this is a mathematics task for which adults have shown sensitivity to spatial layout (Landy & Goldstone, 2007, 2010).

Not all of the evidence was consistent with this interpretation, however. There was a lack of overlap for other tasks with seemingly similar spatial "reading" demands, such as CMAT Charts and Graphs or Geometry. It makes sense, for example, that reading graphs is aided by some facility with perceiving lengths, and visually coordinating bars with scales along axes. One explanation might be that the format of these tests had a verbal component, and this may have obscured their spatial components once verbal ability was controlled. With respect to the CMAT Geometry subtest, a second explanation may have to do with the level of reasoning required. Although the CMAT Geometry subtest included problems that required reasoning, it arguably emphasized procedural knowledge more than the CMAT Algebra or Word Problems subtests, and thus, may not have been as strongly related to spatial skill for that reason. Finally, this finding does not mean there are not processes shared between these tests and spatial skill, but rather, that they did not have particularly strong or unique overlap.

Spatial Scaling

Regarding the third potential type of processing overlap—spatial scaling and magnitude representation—we found no direct evidence in the present study. Specifically, there were no significant cross-domain effects involving either number line estimation or map reading—the two tasks most closely related to spatial scaling in the present battery. Admittedly, map reading is not the purest possible test of spatial scaling as it involves other skills, such as relating analogous parts, and, for some items, shifting perspective. Perhaps a more direct test of spatial scaling, such as Boyer and Levine's (2012) proportional reasoning task, would have less overlap with spatial tasks such as mental rotation or perspective-taking, and thus might provide a cleaner test of the relation of spatial scaling to mathematical skills. Still, it was surprising that a mathematics task such as number line estimation, which has a strong spatial component and seems particularly likely to involve spatial scaling, shared no unique variance with spatial skill as a whole. Perhaps, as other have found, number line estimation is more math-related than spatial (LeFevre et al., 2013).

Developmental Changes

With respect to previous research linking spatial skill and mathematics in childhood, the present results replicated several key findings. First, the relatively strong relation between mental rotation and mathematics performance in kindergarten and third-grade students is consistent with previous correlational research (Carr et al., 2015; Gunderson et al., 2012; Thompson et al., 2013), as well as Cheng and Mix's (2014) training study. Also consistent with previous research, VSWM and VMI had significant loadings with mathematics in third and sixth grade (Laski et al., 2013; Li & Geary, 2013; McKenzie et al., 2003; Sortor & Kulp, 2003). Note that these specific relations held even when other spatial tasks were included and their shared variance was taken into account.

The pattern of an initially greater role for mental rotation in kindergarten, followed by a mixture of weaker relations in third grade, and finally a greater role for VSWM and VMI in sixth grade suggests a gradually shifting relation between spatial skill and mathematics. Perhaps this relation is rooted, early in development, in the underlying spatial representation of numerical relations and operations, thus sharing more processing with mental rotation ability and other visualization tasks, but moves toward more perceptual tracking and memory abilities, such as analyzing the spatial positions of written symbols. The finding of relatively weak relations involving a broad range of spatial tasks in third grade, including mental rotation, block design, perspective-taking, VSWM and VMI, is characteristic of the developmental instability that occurs during transitions (Perry, Church, & Goldin-Meadow, 1988; Schoner, 2008; Siegler, 2007; Thelen et al., 1993), and may reflect a mixture of relations that exist as children shift from one dominant processing mode to the other. This shift could be due to changes in the demands of mathematics content, developmental changes in spatial skill, or both.

For example, the concepts underlying long division are not that different from the concepts underlying simple division, but the perceptual-motor demands of the long division procedure increase in complexity with each additional digit. Perhaps conceptualizing division draws upon spatial representations like those used to imagine objects rotating into different positions, but carrying out complicated multistep mathematics procedures requires the kind of detailed spatial processing that is measured in memory for multiple locations in space or chunking and copying complex figures. Similarly, understanding the rough, holistic value of simple fractions or common multidigit numbers may be achieved through verbal association or spatial scaling in early childhood, but the fine-grained perceptual skills needed to interpret more complex fraction notation in a precise way may require stronger spatial perception and memory.

This pattern may relate more generally to the conceptual-procedural distinction in mathematical development in the sense that some forms of spatial ability could have an important role in conceptual grounding (i.e., spatial visualization for grounding place value meaning), whereas others may come into play when procedural fluency is emphasized (i.e., VSWM for carrying multidigit operations). Consistent with this, recent research has reported that girls who use more sophisticated, decomposition strategies in calculation tasks also have better spatial visualization ability (i.e., using mental rotation and block design measures) than girls who rely on less sophisticated procedural strategies, like counting (Laski et al., 2013).

Indeed, this shift from conceptual grounding to procedural fluency, and a possibly linked shift in associated spatial skills, may be recapitulated several times developmentally as tasks shift from novel to familiar. Recall that when we divided the mathematics items into new and familiar content, some spatial skills related to both, but there were a few cases of a clear, distinct link between spatial skill and old versus new content. For example, in kindergarten, only new content (i.e., mostly multidigit numeracy and operations) was related to spatial visualization tasks, including Block Design and Perspective-Taking, and only familiar content was related to the VMI. In older children, spatial visualization (i.e., Block Design) was related to new content only. This pattern is consistent with previous hypotheses that spatial skills may play a special role when children are first acquiring new concepts (Ackerman, 1988; Uttal & Cohen, 2012), but it goes further by suggesting it is not simply spatial skill that is useful at these early stages, but rather, spatial visualization in particular. For familiar content, as operations become more automatic and retrievable, spatial visualization may play less of a role, but other spatial skills may increase in importance, as we saw in the kindergarten effects with VMI for familiar content only and VMI/VSWM for sixth grade overall.

The age-related patterns we obtained are intriguing, but they are also preliminary. Further study is needed to confirm these shifts in other populations, using alternative tasks, and in individual children using a longitudinal design. Also, whereas our data can suggest relations, they do not establish causality or even provide direct evidence for the shared processes we have posited in our interpretations. Future research using training may be one way to test these ideas more directly.

Conclusions

The present study indicates that spatial skill and mathematics are separate but broadly overlapping domains with a few skills that exhibit particularly strong cross-domain linkages. This general overlap does not change from kindergarten to sixth grade, but the specific linkages do, and the specific linkages also shift depending upon the familiarity of the mathematics content.

These results may have educational implications but further research with training designs is needed. There has been a great deal of interest in the effects of cognitive training on academic outcomes generally speaking (see Schubert, Strobach, & Karbach, 2014, for an overview), and calls for such research in the area of spatial cognition in particular (Lubinski, 2010; Levine, Foley, Lourenco, Ehrlich, & Ratliff, 2016; NCTM, 2010; Newcombe, 2010, 2013; Sorby, 2009; Uttal et al., 2013; Verdine, Irwin, Golinkoff, & Hirsh-Pasek, 2014). However, initial attempts to show transfer of domain general training to academic achievement have yielded mixed results, with some studies demonstrating transfer of working memory and executive function to academic tasks (e.g., Alloway, 2012; Dahlin, 2011) and others failing to do so (e.g., Holmes, Gathercole, & Dunning, 2009; see Melby-Lervåg & Hulme, 2013 and Titz & Karbach, 2014, for reviews). This has been the case as well for the few spatial training studies involving mathematics (i.e., Cheng & Mix, 2014 vs. Hawes, Moss, Caswell, & Poliszczuk, 2015).

A key issue in reconciling these discrepancies may be the degree of shared processing among various subskills. As others have pointed out, the crux of cognitive training is not to improve performance on a particular task, but to improve performance of a latent ability that is expressed in multiple tasks (Noack, Lovden, & Schmiedek, 2014). Our results are a step toward understanding what those latent abilities might be for spatial skill and mathematics. Only training experiments can determine whether these latent abilities are malleable and whether this training generalizes to academic learning. Indeed, a pertinent question is whether such training, if effective, offers greater benefit than simply increasing instruction in the academic skill itself. However, by starting with a clearer understanding of what these latent abilities are, researchers have the greatest chance of leveraging them effectively.

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. http://dx.doi.org/10.1037/0096-3445.117.3.288

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131,* 30–60. http://dx.doi.org/10.1037/0033-2909.131.1.30

Alloway, T. P. (2012). Can interactive working memory training improve learning? *Journal of Interactive Learning Research, 23,* 197–207.

Alloway, T. P., & Passolunghi, M. C. (2011). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences, 21,* 133–137. http://dx.doi.org/10.1016/j.lindif.2010.09.013

Barakat, M. K. (1951). A factorial study of mathematical abilities. *British Journal of Psychology, 4,* 137–156.

Baroody, A. J., Feil, K., & Johnson, A. R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education, 38,* 115–131.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59,* 617–645. http://dx.doi.org/10.1146/annurev.psych.59.103006.093639

Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental*

Science, 14, 125–135. http://dx.doi.org/10.1111/j.1467-7687.2010.00962.x

Beery, K. E., & Beery, N. A. (2010). *The Beery–Buktenica developmental test of visual-motor integration: Beery VMI with supplemental developmental tests of visual perception and motor coordination: Administration, scoring and teaching manual* (6th ed.). Minneapolis, MN: Pearson.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42,* 189–201. http://dx.doi.org/10.1037/0012-1649.41.6.189

Boyer, T. W., & Levine, S. C. (2012). Child proportional scaling: Is 1/3 = 2/6 = 3/9 = 4/12? *Journal of Experimental Child Psychology, 111,* 516–533. http://dx.doi.org/10.1016/j.jecp.2011.11.001

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–139). Newbury Park, CA: Sage.

Burgess, N. (2006). Spatial memory: How egocentric and allocentric combine. *Trends in Cognitive Sciences, 10,* 551–557. http://dx.doi.org/10.1016/j.tics.2006.10.005

Burnett, S. A., Lane, D. M., & Dratt, L. M. (1979). Spatial visualization and sex differences in quantitative ability. *Intelligence, 3,* 345–354. http://dx.doi.org/10.1016/0160-2896(79)90003-5

Carr, M., Alexeev, N., Horan, E., Barned, N., & Wang, L. (2015, March). *Spatial skills in elementary school predict trajectory of number sense development and achievement*. Poster presented at the biennial meeting of the Society for Research in Child Development, Philadelphia, PA.

Carr, M., Steiner, H. H., Kyser, B., & Biddlecomb, B. (2008). A comparison of predictors of early emerging gender differences in mathematics competency. *Learning and Individual Differences, 18,* 61–75. http://dx.doi.org/10.1016/j.lindif.2007.04.005

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, England: University of Cambridge Press. http://dx.doi.org/10.1017/CBO9780511571312

Casey, M. B., Nuttall, R. L., & Pezaris, E. (2001). Spatial-mechanical reasoning skills versus mathematical self-confidence as mediators of gender differences on mathematics subtests using cross-national gender-based items. *Journal for Research in Mathematics Education, 32,* 28–57. http://dx.doi.org/10.2307/749620

Casey, M. B., Nuttall, R., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology, 31,* 697–705. http://dx.doi.org/10.1037/0012-1649.31.4.697

Cauley, K. M. (1988). Construction of logical knowledge: Studying of borrowing in subtraction. *Journal of Educational Psychology, 80,* 202–205. http://dx.doi.org/10.1037/0022-0663.80.2.202

Caviola, S., Mammarella, I. C., Cornoldi, C., & Lucangeli, D. (2012). The involvement of working memory in children's exact and approximate mental addition. *Journal of Experimental Child Psychology, 112,* 141–160. http://dx.doi.org/10.1016/j.jecp.2012.02.005

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods & Research, 36,* 462–494. http://dx.doi.org/10.1177/0049124108314720

Cheng, Y.-L., & Mix, K. S. (2014). Spatial training improves children's mathematics ability. *Journal of Cognition and Development, 15,* 2–11. http://dx.doi.org/10.1080/15248372.2012.725186

Chou, C. P., & Bentler, P. M. (1995). Estimation and tests in structural equation modeling. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 37–55). Thousand Oaks, CA: Sage.

Chou, C. P., Bentler, P. M., & Satorra, A. (1991). Scaled test statistics and robust standard errors for non-normal data in covariance structure analysis: A Monte Carlo study. *The British Journal of Mathematical and Statistical Psychology, 44,* 347–357. http://dx.doi.org/10.1111/j.2044-8317.1991.tb00966.x

Cobb, P., & Wheatley, G. (1988). Children's initial understandings of ten. *Focus on Learning Problems in Mathematics, 10,* 1–28.

Colom, R., Contreras, M. J., Botella, J., & Santacreu, J. (2002). Vehicles of spatial ability. *Personality and Individual Differences, 32,* 903–912. http://dx.doi.org/10.1016/S0191-8869(01)00095-2

Cornell, E. H., & Hay, D. H. (1984). Children's acquisition of a route via different media. *Environment and Behavior, 16,* 627–641. http://dx.doi.org/10.1177/0013916584165005

Cowey, A., Small, M., & Ellis, S. (1994). Left visuo-spatial neglect can be worse in far than in near space. *Neuropsychologia, 32,* 1059–1066. http://dx.doi.org/10.1016/0028-3932(94)90152-X

Cudeck, R., & O'Dell, L. L. (1994). Applications of standard error estimates in unrestricted factor analysis: Significance tests for factor loadings and correlations. *Psychological Bulletin, 115,* 475–487. http://dx.doi.org/10.1037/0033-2909.115.3.475

Dahlin, E. (2011). Effects of working memory training on reading in children with special needs. *Reading and Writing, 24,* 479–491. http://dx.doi.org/10.1007/s11145-010-9238-y

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122,* 371–396. http://dx.doi.org/10.1037/0096-3445.122.3.371

Delgado, A. R., & Prieto, G. (2004). Cognitive mediators and sex-related differences in mathematics. *Intelligence, 32,* 25–32. http://dx.doi.org/10.1016/S0160-2896(03)00061-8

Dulaney, A. (2014). *The roles of visuospatial and verbal working memory in children's mathematical performance* (Doctoral dissertation). Available from ProQuest Dissertations & Theses Global. (Order No. 3615784)

Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology, 123,* 53–72. http://dx.doi.org/10.1016/j.jecp.2014.01.013

Fouracre, L. (1926). Psychology tests of mathematical ability. *Forum of Education, 3,* 201–205.

Freedman, D. A. (2006). On the so-called "Huber Sandwich Estimator" and "robust standard errors." *The American Statistician, 60,* 299–302. http://dx.doi.org/10.1198/000313006X152207

Frick, A., Möhring, W., & Newcombe, N. S. (2014). Picturing perspectives: Development of perspective-taking abilities in 4- to 8-year-olds. *Frontiers in Psychology, 5,* 386. http://dx.doi.org/10.3389/fpsyg.2014.00386

Fuson, K. C., & Briars, D. J. (1990). Using a base-ten blocks learning/teaching approach for first- and second-grade place-value and multidigit addition and subtraction. *Journal for Research in Mathematics Education, 21,* 180–206. http://dx.doi.org/10.2307/749373

Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *The British Journal of Educational Psychology, 70,* 177–194. http://dx.doi.org/10.1348/000709900158047

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114,* 345–362. http://dx.doi.org/10.1037/0033-2909.114.2.345

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47,* 1539–1552. http://dx.doi.org/10.1037/a0025510

Geary, D. C. (2013). Early foundations for mathematics learning and their relations to learning disabilities. *Current Directions in Psychological Science, 22,* 23–27. http://dx.doi.org/10.1177/0963721412469398

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development, 78,* 1343–1359. http://dx.doi.org/10.1111/j.1467-8624.2007.01069.x

Geary, D. C., Saults, S. J., Liu, F., & Hoard, M. K. (2000). Sex differences in spatial cognition, computational fluency, and arithmetical reasoning. *Journal of Experimental Child Psychology, 77,* 337–353. http://dx.doi.org/10.1006/jecp.2000.2594

Ginsburg, H. P., & Baroody, A. J. (2003). *Test of Early Mathematics Ability manual* (3rd ed.). Austin, TX: Pro-Ed.

Glenberg, A. M., Brown, M., & Levin, J. R. (2007). Enhancing comprehension in small reading groups using a manipulation strategy. *Contemporary Educational Psychology, 32,* 389–399. http://dx.doi.org/10.1016/j.cedpsych.2006.03.001

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,* 1229–1241. http://dx.doi.org/10.1037/a0027433

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology, 44,* 1457–1465. http://dx.doi.org/10.1037/a0012682

Hambrick, Z. D., Kane, J. M., & Engle, W. R. (2005). The role of working memory in higher-level cognition. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and intelligence: Identifying the mechanisms of the mind* (pp. 104–121). New York, NY: Cambridge University Press.

Hawes, Z., Moss, J., Caswell, B., & Poliszczuk, D. (2015). Effects of mental rotation training on children's spatial and mathematical performance: A randomized controlled study. *Trends in Neuroscience and Education, 4,* 60–68. http://dx.doi.org/10.1016/j.tine.2015.05.001

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence, 32,* 175–191. http://dx.doi.org/10.1016/j.intell.2003.12.001

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Hillsdale, NJ: Erlbaum.

Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations—A meta-analytic review. *Educational Psychology Review, 22,* 245–269. http://dx.doi.org/10.1007/s10648-010-9126-7

Holmes, J., Adams, J. W., & Hamilton, C. J. (2008). The relationship between visuospatial sketchpad capacity and children's mathematical skills. *The European Journal of Cognitive Psychology, 20,* 272–289. http://dx.doi.org/10.1080/09541440701612702

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science, 12,* F9–F15. http://dx.doi.org/10.1111/j.1467-7687.2009.00848.x

Holzinger, K. J., & Harman, H. H. (1938). Comparison of two factorial analyses. *Psychometrika, 3,* 45–60. http://dx.doi.org/10.1007/BF02287919

Holzinger, K. J., & Swineford, F. (1946). The relation of two bi-factors to achievement in geometry and other subjects. *Journal of Educational Psychology, 37,* 257–265. http://dx.doi.org/10.1037/h0053536

Hresko, W., Schlieve, P., Herron, S., Swain, C., & Sherbenou, R. (2003). *Comprehensive Mathematical Abilities Test (CMAT).* Austin: PRO-ED.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. http://dx.doi.org/10.1080/10705519909540118

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience, 6,* 435–448. http://dx.doi.org/10.1038/nrn1684

Huttenlocher, J., Jordan, N. C., & Levine, S. C. (1994). A mental model for early arithmetic. *Journal of Experimental Psychology: General, 123,* 284.

Huttenlocher, J., & Presson, C. C. (1973). Mental rotation and the perspective problem. *Cognitive Psychology, 4,* 277–299. http://dx.doi.org/10.1016/0010-0285(73)90015-7

Jesson, D. S. J. (1983). The development of place value skills in primary and middle school children. *Research in Education, 29,* 69–79.

Johnson, E. S. (1998). An exploration of the relation between mathematics achievement and factors of intelligence (Doctoral dissertation). *Dissertation Abstracts International: Section A. The Humanities and Social Sciences, 57*(11), 4692.

Kamii, C. (1986). Place value: An explanation of its difficulty and educational implications for the primary grades. *Journal of Research in Childhood Education, 1,* 75–86. http://dx.doi.org/10.1080/02568548609594909

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.

Kelley, T. I. (1928). *Crossroads in the mind of man; A study of differentiable mental abilities*. Stanford, CA: Stanford University Press. http://dx.doi.org/10.1037/11206-000

Kesner, R. P., Farnsworth, G., & DiMattia, B. V. (1989). Double dissociation of egocentric and allocentric space following medial prefrontal and parietal cortex lesions in the rat. *Behavioral Neuroscience, 103,* 956–961. http://dx.doi.org/10.1037/0735-7044.103.5.956

Kline, R. B. (2005). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education, 37,* 297–312.

Kosslyn, S. M., Koenig, O., Barrett, A., Cave, C. B., Tang, J., & Gabrieli, J. D. E. (1989). Evidence for two types of spatial representations: Hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 723–735. http://dx.doi.org/10.1037/0096-1523.15.4.723

Kouba, V. L., Brown, C. A., Carpenter, T. P., Lindquist, M. M., Silver, E. A., & Swafford, J. O. (1988). Results of the fourth NAEP assessment of mathematics: Number, operations, and world problems. *The Arithmetic Teacher, 35,* 14–19.

Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition, 29,* 745–756. http://dx.doi.org/10.3758/BF03200477

Krajewski, K., & Schneider, W. (2009). Exploring the impact of phonological awareness, visual-spatial working memory, and preschool quantity-number competencies on mathematics achievement in elementary school: Findings from a 3-year longitudinal study. *Journal of Experimental Child Psychology, 103,* 516–531. http://dx.doi.org/10.1016/j.jecp.2009.03.009

Kyttälä, M., Aunio, P., Lehto, J. E., Van Luit, J., & Hautamaki, J. (2003). Visuospatial working memory and early numeracy. *Educational and Child Psychology, 20,* 65–76.

Labinowicz, E. (1985). *Learning from children: New beginnings for teaching numerical thinking*. Menlo Park, CA: Addison Wesley.

Lakoff, G., & Núñez, R. (2000). *Where mathematics comes from: How the embodies brings mathematics into being*. New York, NY: Basic Books.

Landy, D., & Goldstone, R. L. (2007). The alignment of ordering and space in arithmetic computation. In *Proceedings of the 29th Annual Meeting of the Cognitive Science Society* (pp. 437–442). Philadelphia, PA: Cognitive Science Society.

Landy, D., & Goldstone, R. L. (2010). Proximity and precedence in arithmetic. *Quarterly Journal of Experimental Psychology, 63,* 1953–1968. http://dx.doi.org/10.1080/17470211003787619

Larsen, R., & Warne, R. T. (2010). Estimating confidence intervals for eigenvalues in exploratory factor analysis. *Behavior Research Methods, 42,* 871–876. http://dx.doi.org/10.3758/BRM.42.3.871

Laski, E. V., Casey, B. M., Yu, Q., Dulaney, A., Heyman, M., & Dearing, E. (2013). Spatial skills as a predictor of first grade girls' use of higher level arithmetic strategies. *Learning and Individual Differences, 23,* 123–130. http://dx.doi.org/10.1016/j.lindif.2012.08.001

LeFevre, J.-A., Fast, L., Skwarchuk, S. L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development, 81,* 1753–1767. http://dx.doi.org/10.1111/j.1467-8624.2010.01508.x

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. http://dx.doi.org/10.3389/fpsyg.2013.00641

Levine, S. C., Foley, A., Lourenco, S., Ehrlich, S., & Ratliff, K. (2016). Sex differences in spatial cognition: Advancing the conversation. *Wiley Interdisciplinary Reviews: Cognitive Science, 7,* 127–155. http://dx.doi.org/10.1002/wcs.1380

Li, Y., & Geary, D. C. (2013). Developmental gains in visuospatial memory predict gains in mathematics achievement. *PLoS ONE, 8,* e70160. http://dx.doi.org/10.1371/journal.pone.0070160

Liben, L. S. (2001). Thinking through maps. In M. Gattis (Ed.), *Spatial schemas and abstract thought* (pp. 45–77). Cambridge, MA: MIT Press.

Liben, L. S., & Downs, R. M. (1989). Understanding maps as symbols: The development of map concepts in children. *Advances in Child Development and Behavior, 22,* 145–201. http://dx.doi.org/10.1016/S0065-2407(08)60414-0

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56,* 1479–1498. http://dx.doi.org/10.2307/1130467

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition, 22,* 395–410. http://dx.doi.org/10.3758/BF03200866

Lohman, D. F. (1979). *Spatial ability: A review and reanalysis of the correlational literature* (Tech. Rep. No. 8). Stanford, CA: Aptitude Research Project, School of Education, Stanford University.

Lohman, D. F. (1988). Spatial abilities as traits, processes, and knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 4, pp. 181–248). Hillsdale, NJ: Erlbaum, Inc.

Lohman, D. F. (1996). Spatial ability and g. In I. Dennis & P. Tapsfield (Eds.), *Human abilities: Their nature and measurement* (pp. 97–116). Hillsdale, NJ: Erlbaum.

Lourenco, S. F., & Longo, M. R. (2009). The plasticity of near space: Evidence for contraction. *Cognition, 112,* 451–456. http://dx.doi.org/10.1016/j.cognition.2009.05.011

Lovegrove, W., Martin, F., Bowling, A., Blackwood, M., Badcock, D., & Paxton, S. (1982). Contrast sensitivity functions and specific reading disability. *Neuropsychologia, 20,* 309–315. http://dx.doi.org/10.1016/0028-3932(82)90105-1

Lubinski, D. (2010). Spatial ability and STEM: A sleeping giant for talent identification and development. *Personality and Individual Differences, 49,* 344–351. http://dx.doi.org/10.1016/j.paid.2010.03.022

Lubinski, D., & Benbow, C. P. (1992). Gender differences in abilities and preferences among the gifted: Implications for the math-science pipeline. *Current Directions in Psychological Science, 1,* 61–66. http://dx.doi.org/10.1111/1467-8721.ep11509746

Markey, S. M. (2010). The relationship between visual-spatial reasoning ability and math and geometry problem-solving. *Dissertation Abstracts International: Section B*. The Sciences and Engineering, *59,* 724.

Marzolf, D. P., & DeLoache, J. S. (1994). Transfer in young children's understanding of spatial representations. *Child Development, 65,* 1–15. http://dx.doi.org/10.2307/1131361

Matthews, P. G., Chesney, D. L., & McNeil, N. M. (2014). Are fractions natural numbers too? In M. P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 982–987). Austin TX: Cognitive Science Society.

Mazzocco, M. M. M., & Myers, G. F. (2003). Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Annals of Dyslexia, 53,* 218–253. http://dx.doi.org/10.1007/s11881-003-0011-7

McKenzie, B., Bull, R., & Gray, C. (2003). The effects of phonological and visual-spatial interference on children's arithmetical performance. *Educational and Child Psychology, 20,* 93–108.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49,* 270–291. http://dx.doi.org/10.1037/a0028228

Meyer, M. L., Salimpoor, V. N., Wu, S. S., Geary, D. C., & Menon, V. (2010). Differential contribution of specific working memory components to mathematics achievement in 2nd and 3rd graders. *Learning and Individual Differences, 20,* 101–109. http://dx.doi.org/10.1016/j.lindif.2009.08.004

Michael, B. W., Guilford, P. J., Fruchter, B., & Zimmerman, S. W. (1957). The description of spatial-visualization abilities. *Educational and Psychological Measurement, 17,* 185–199. http://dx.doi.org/10.1177/001316445701700202

Mix, K. S., & Cheng, Y.-L. (2012). Space and math: The development and educational implications. In J. Benson (Ed.), *Advances in child development and behavior* (pp. 179–243). New York, NY: Elsevier.

Mix, K. S., Prather, R. W., Smith, L. B., & Stockton, J. D. (2014). Young children's interpretation of multidigit number names: From emerging competence to mastery. *Child Development, 85,* 1306–1319. http://dx.doi.org/10.1111/cdev.12197

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,* 621–640. http://dx.doi.org/10.1037/0096-3445.130.4.621

Moeller, K., Pixner, S., Zuber, J., Kaufmann, L., & Nuerk, H. C. (2011). Early place-value understanding as a precursor for later arithmetic performance—A longitudinal study on numerical development. *Research in Developmental Disabilities, 32,* 1837–1851. http://dx.doi.org/10.1016/j.ridd.2011.03.012

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and Standards for school mathematics*. Reston, VA: Author.

National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common Core state standards for mathematics*. Washington, DC: Author.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353–383. http://dx.doi.org/10.1016/0010-0285(77)90012-3

Neuburger, S., Jansen, P., Heil, M., & Quaiser-Pohl, C. (2011). Gender differences in pre-adolescents' mental-rotation performance: Do they depend on grade and stimulus type? *Personality and Individual Differences, 50,* 1238–1242. http://dx.doi.org/10.1016/j.paid.2011.02.017

Newcombe, N. S. (2010). Picture this: Increasing math and science learning by improving spatial thinking. *American Educator, 34,* 29–35, 43.

Newcombe, N. S. (2013). Seeing relationships: Using spatial thinking to teach science, mathematics, and social studies. *American Educator, 37,* 26–31, 40.

Newcombe, N. S., & Huttenlocher, J. (1992). Children's early ability to solve perspective-taking problems. *Developmental Psychology, 28,* 635–643. http://dx.doi.org/10.1037/0012-1649.28.4.635

Newcombe, N. S., Levine, S. C., & Mix, K. S. (2016). Thinking about quantity: The intertwined development of spatial and numerical cognition. *Wiley Interdisciplinary Reviews: Cognitive Science, 6,* 491–505. http://dx.doi.org/10.1002/wcs.1369

Newcombe, N. S., & Shipley, T. F. (2015). Thinking about spatial thinking: New typology, new assessments. In J. S. Gero (Ed.), *Studying visual and spatial reasoning for design creativity* (pp. 179–192). New York, NY: Springer. http://dx.doi.org/10.1007/978-94-017-9297-4_10

Noack, H., Lövdén, M., & Schmiedek, F. (2014). On the validity and generality of transfer effects in cognitive training research. *Psychological Research, 78,* 773–789. http://dx.doi.org/10.1007/s00426-014-0564-6

Novack, M., Brooks, N., Kennedy, D., Levine, S., & Goldin-Meadow, S. (October, 2013). *Using Action and Gesture to Improve Mental Rotation*. Poster presented at the Cognitive Development Society Conference, Memphis, TN.

Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology, 55,* 169–195.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3,* 359–400. http://dx.doi.org/10.1016/0885-2014(88)90021-4

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test: Different versions and factors that affect performance. *Brain and Cognition, 28,* 39–58. http://dx.doi.org/10.1006/brcg.1995.1032

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20,* 110–122. http://dx.doi.org/10.1016/j.lindif.2009.10.005

Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of Experimental Child Psychology, 91,* 137–157. http://dx.doi.org/10.1016/j.jecp.2005.01.004

Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling* (2nd ed.). Mahwah, NJ: Erlbaum.

Raykov, T., & Marcoulides, G. A. (2008). *An introduction to applied multivariate analysis*. New York, NY: Taylor & Francis.

Resnick, L. B., & Omanson, S. F. (1987). Learning to understand arithmetic. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 41–95). Hillsdale, NJ: Erlbaum.

Rieser, J. J., Garing, A. E., & Young, M. F. (1994). Imagery, action, and young children's spatial orientation: It's not being there that counts, it's what one has in mind. *Child Development, 65,* 1262–1278. http://dx.doi.org/10.2307/1131498

Robert, N. D., & LeFevre, J.-A. (2013). Ending up with less: The role of working memory in solving simple subtraction problems with positive and negative answers. *Research in Mathematics Education, 15,* 165–176. http://dx.doi.org/10.1080/14794802.2013.797748

Robinson, N., Abbott, R., Berninger, V. W., & Busse, J. (1996). Structure of abilities in math-precocious young children: Gender similarities and differences. *Journal of Educational Psychology, 88,* 341–352. http://dx.doi.org/10.1037/0022-0663.88.2.341

Sattler, J. M. (2001). *Assessment of children: Cognitive applications* (4th ed.). San Diego, CA: Author.

Schmitt, T. A., & Sass, D. A. (2011). Rotation criteria and hypothesis testing for exploratory factor analysis: Implications for factor pattern loadings and interfactor correlations. *Educational and Psychological Measurement, 71,* 95–113. http://dx.doi.org/10.1177/0013164410387348

Schneider, M., Grabner, R. H., & Paetsch, J. (2009). Mental number line, number line estimation, and mathematical achievement: Their interrelations in Grades 5 and 6. *Journal of Educational Psychology, 101,* 359–372. http://dx.doi.org/10.1037/a0013840

Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology, 47,* 1525–1538. http://dx.doi.org/10.1037/a0024997

Schöner, G. (2008). Dynamical systems approaches to cognition. In R. Sun (Ed.), *Cambridge handbook of computational cognitive modeling* (pp. 101–126). New York, NY: Cambridge University Press.

Schubert, T., Strobach, T., & Karbach, J. (2014). New directions in cognitive training: On methods, transfer, and application. *Psychological Research, 78,* 749–755. http://dx.doi.org/10.1007/s00426-014-0619-8

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125,* 4–27. http://dx.doi.org/10.1037/0096-3445.125.1.4

Shusterman, A., Ah Lee, S., & Spelke, E. S. (2008). Young children's spontaneous use of geometry in maps. *Developmental Science, 11,* F1–F7. http://dx.doi.org/10.1111/j.1467-7687.2007.00670.x

Siegler, R. S. (2007). Cognitive variability. *Developmental Science, 10,* 104–109. http://dx.doi.org/10.1111/j.1467-7687.2007.00571.x

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child development, 75,* 428–444.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14,* 237–250. http://dx.doi.org/10.1111/1467-9280.02438

Slusser, E. B., Santiago, R. T., & Barth, H. C. (2013). Developmental change in numerical estimation. *Journal of Experimental Psychology: General, 142,* 193–208. http://dx.doi.org/10.1037/a0028560

Smith, I. M. (1964). *Spatial ability: Its educational and social significance*. San Diego, CA: Knapp.

Sorby, S. A. (2009). Educational research in developing 3-D spatial skills for engineering students. *International Journal of Science Education, 31,* 459–480. http://dx.doi.org/10.1080/09500690802595839

Sortor, J. M., & Kulp, M. T. (2003). Are the results of the Beery-Buktenica Developmental Tests of Visual-Motor Integration related to achievement test scores? *Optometry and Vision Science, 80,* 758–763.

Spearman, C., & Jones, L. (1950). *Human ability*. Oxford, England: Macmillan.

Stafylidou, S., & Vosniadou, S. (2004). The development of students' understanding of the numerical value of fractions. *Learning and Instruction, 14,* 503–518.

Star, J. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education, 36,* 404–411.

Steiger, J. H. (1989). *EzPATH: A supplementary module for SYSTAT and SYGRAPH*. Evanston, IL: SYSTAT.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25,* 173–180. http://dx.doi.org/10.1207/s15327906mbr2502_4

Stieff, M. (2013). Sex differences in the mental rotation of chemistry representations. *Journal of Chemical Education, 90,* 165–170. http://dx.doi.org/10.1021/ed300499t

Thelen, E., Corbetta, D., Kamm, K., Spencer, J. P., Schneider, K., & Zernicke, R. F. (1993). The transition to reaching: Mapping intention and intrinsic dynamics. *Child Development, 64,* 1058–1098. http://dx.doi.org/10.2307/1131327

Thompson, J. M., Nuerk, H.-C., Moeller, K., & Cohen Kadosh, R. (2013). The link between mental rotation ability and basic numerical representations. *Acta Psychologica, 144,* 324–331. http://dx.doi.org/10.1016/j.actpsy.2013.05.009

Thurstone, L. L. (1944). *A factorial study of perception*. Chicago, IL: University of Chicago Press.

Titz, C., & Karbach, J. (2014). Working memory and executive functions: Effects of training on academic achievement. *Psychological Research, 78,* 852–868. http://dx.doi.org/10.1007/s00426-013-0537-1

Trbovich, P. L., & LeFevre, J.-A. (2003). Phonological and visual working memory in mental addition. *Memory & Cognition, 31,* 738–745. http://dx.doi.org/10.3758/BF03196112

Uttal, D. H. (2000). Seeing the big picture: Map use and the development of spatial cognition. *Developmental Science, 3,* 247–264. http://dx.doi.org/10.1111/1467-7687.00119

Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why and how. *Psychology of Learning and Motivation, 57,* 147–181. http://dx.doi.org/10.1016/B978-0-12-394293-7.00004-2

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin, 139,* 352–402. http://dx.doi.org/10.1037/a0028446

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47,* 599–604. http://dx.doi.org/10.2466/pms.1978.47.2.599

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. http://dx.doi.org/10.1016/j.jecp.2014.02.012

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117,* 250–270. http://dx.doi.org/10.1037/0033-2909.117.2.250

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences, 7,* 483–488. http://dx.doi.org/10.1016/j.tics.2003.09.002

Wang, J., & Wang, X. (2012). *Structural equation modeling: Applications using Mplus*. Hoboken, NJ: Wiley. http://dx.doi.org/10.1002/9781118356258

Wechsler, D., Kaplan, E., Fein, D., Kramer, J., Morris, R., Delis, D., & Maerlender, A. (2004). *WISC–IV: Wechsler Intelligence Scale for Children integrated technical and interpretive manual* (4th ed.). Minneapolis, MN: Pearson.

Werdelin, I. (1958). *The mathematical ability: Experimental and factorial studies*. Oxford, England: Gleerup.

Wilson, J. H. (1933). Group factors among abilities involved in a school certificate examination. *The British Journal of Educational Psychology, 3,* 71–86. http://dx.doi.org/10.1111/j.2044-8279.1933.tb02904.x

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson tests of achievement*. Itasca, IL: Riverside.

Wrigley, J. (1958). The factorial nature of ability in elementary mathematics. *The British Journal of Educational Psychology, 28,* 61–78. http://dx.doi.org/10.1111/j.2044-8279.1958.tb01426.x

Young, C. J., & Opfer, J. E. (2011). Psychophysics of number representation: A unified approach to single and multi-digit magnitude estimation. *Zeitschrift für Psychologie/The Journal of Psychology, 219,* 58–63.

Received May 4, 2015
Revision received April 13, 2016
Accepted April 22, 2016 ■