Journal of Experimental Child Psychology 145 (2016) 95–119

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Cognitive and numerosity predictors of mathematical skills in middle school

Paul T. Cirino [a,*], Tammy D. Tolar [a], Lynn S. Fuchs [b], Emily Huston-Warren [a]

[a] University of Houston, Houston, TX 77204, USA
[b] Vanderbilt University, Nashville, TN 37235, USA

ARTICLE INFO

Article history:
Received 29 September 2015
Revised 18 December 2015

Keywords:
Working memory
Language
Spatial
Numerosity
Fractions
Proportional reasoning
Mathematics
Middle school

ABSTRACT

There is a strong research base on the underlying concomitants of early developing math skills. Fewer studies have focused on later developing skills. Here, we focused on direct and indirect contributions of cognitive measures (e.g., language, spatial skills, working memory) and numerosity measures, as well as arithmetic proficiency, on key outcomes of fraction performance, proportional reasoning, and broad mathematics achievement at sixth grade ($N = 162$) via path analysis. We expected a hierarchy of skill development, with predominantly indirect effects of cognitive factors via number and arithmetic. Results controlling for age showed that the combination of cognitive, number, and arithmetic variables cumulatively accounted for 38% to 44% of the variance in fractions, proportional reasoning, and broad mathematics. There was consistency across outcomes, with more proximal skills providing direct effects and with the effects of cognitive skills being mediated by number and by more proximal skills. Results support a hierarchical progression from domain-general cognitive processes through numerosity and arithmetic skills to proportional reasoning to broad mathematics achievement.

* Corresponding author.
E-mail address: pcirino@uh.edu (P.T. Cirino).

http://dx.doi.org/10.1016/j.jecp.2015.12.010

Introduction

Mathematical skills are important for successful living in modern society (Bynner, 2012; Gates, 2009; Geary, 2013; Kilpatrick, Swafford, & Findell, 2001; Parsons & Bynner, 2008; Rakes, Valentine, McGatha, & Ronau, 2010; Rose & Betts, 2004; Vogel, 2008). The past 20 years have seen a rapidly expanding literature on the predictors of mathematical skills across ages and types of skill (e.g., Aunio & Niemivirta, 2010; Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Cirino, 2011; Cirino, Morris, & Morris, 2002; Fuchs, Geary, Fuchs, Compton, & Hamlett, 2014; Fuchs et al., 2005, 2008, 2010a, 2010b; Geary, 1993; Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Halberda, Mazzocco, & Feigenson, 2008; Jordan, Hanich, & Kaplan, 2003; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Koponen, Aunola, Ahonen, & Nurmi, 2007; Krajewski & Schneider, 2009; Mazzocco, Feigenson, & Halberda, 2011; Siegler & Ramani, 2009; Siegler et al., 2012). Much of that work has focused on the prediction of early developing mathematical skills (through Grade 3) and has been spurred by the development of influential models relevant to this stage of development (e.g., Butterworth, 1999, 2005; Dehaene, 2001; Dehaene & Cohen, 1995; Feigenson, Dehaene, & Spelke, 2004; Geary, 1993, 2004, 2013; LeFevre et al., 2010; von Aster & Shalev, 2007). These empirical and theoretical contributions have yielded strong insights into the relative contribution of differing types of predictors for the development of primary-grade math skill. Such studies set the stage for research on later developing mathematical skills given the hierarchical nature of mathematics skill learning (e.g., Fuchs et al., 2006; National Mathematics Advisory Panel [NMAP], 2008).

Although few studies have focused on more advanced mathematical skills, there have been notable advances concerning the predictors of fraction performance at later elementary grades (e.g., Hecht & Vagi, 2010; Jordan et al., 2013; Vukovic et al., 2014). The current study sought to contribute to this growing literature by clarifying the roles of individual differences from most distal to most proximal, including cognitive processes (e.g., working memory, language), numerosity (magnitude and estimation), and whole number arithmetic skills, for middle school mathematical competencies. These include fraction performance and proportional reasoning, which are themselves critical influences for later developing skills such as algebra (Hecht, Vagi, & Torgesen, 2007; NMAP, 2008; Siegler et al., 2010). Brown and Quinn (2007a, 2007b) described multiple ways in which algebra critically depends on fraction concepts (e.g., the importance of understanding of slope, a fraction, for linear equations).

Supporting competencies of early math skills

The literature on early math prediction supports contributions of both domain-specific and more general cognitive processes. Domain-specific components typically include magnitude comparison (approximate number system acuity) and/or estimation, and these are potentially foundational to later mathematical outcomes (e.g., Butterworth, 2005; Dehaene & Cohen, 1995; Dehaene, Piazza, Pinel, & Cohen, 2003; Feigenson et al., 2004). These are often collectively referred to as numerosity and are typically distinguished from addition and other arithmetic operations. Mathematics performance also builds on preexisting and developing cognitive processes, including linguistic and spatial skills and especially working memory (e.g., Raghubar, Barnes, & Hecht, 2010). Two widely used models, LeFevre and colleagues (2010) and von Aster and Shalev (2007), emphasize the interaction between domain-general and domain-specific processes for arithmetic but are less clear about their relative contributions to more advanced mathematical skill in the context of more elementary arithmetic. Data bearing on this question often come from the early elementary school grades, when math facts are less likely to be automatized (Ashcraft, 1992; Ashcraft & Christy, 1995; Calhoon, Emerson, Flores, & Houchins, 2007), although it is recognized that this automization is conditional on practice and may vary with age or cultural factors (Geary, Bow-Thomas, Liu, & Siegler, 1996),

Empirical results vary with age and the type of mathematical performance considered, but proficiency in both domain-specific and domain-general areas contribute to outcomes, with the former being more predictive of the earlier developing skills (e.g., basic facts) than later developing or more complex skills (e.g., procedural computation, word problems); in contrast, domain-general processes contribute across the skill range. Typically, more complex mathematical skills have a wider range of unique predictors that together account for more variance in these skills relative to more basic skills (e.g., Cowan & Powell, 2014; Fuchs et al., 2010a, 2010b). For example, Fuchs et al. (2010a) found that both specific processes (e.g., number line estimation, fluency in processing numerical magnitudes with both symbolic and nonsymbolic stimuli) and general processes (e.g., working memory, attention) are relevant to varying degrees for different mathematical skills. However, for first-grade computations, numerosity predictors were more predictive; for first-grade word problems, which are more complex, general factors were more predictive. In a separate study, Fuchs et al. (2010) showed that for computational skill, specific predictors were more relevant than general ones; in contrast, for word problem solving, both types of predictors were equally relevant, with more overall variance being accounted for, and with more unique predictors, than was the case for computations.

Distinctions are present within both domain-specific and domain-general predictors. For example, within numerosity, there are mixed findings regarding whether symbolic or nonsymbolic predictors are more relevant. In students with math learning difficulties, some studies have implicated nonsymbolic magnitude (e.g., Halberda et al., 2008; Mazzocco et al., 2011), others have implicated symbolic magnitude (e.g., Cirino, 2011; Holloway & Ansari, 2009; Lyons, Ansari, & Beilock, 2012; Mundy & Gilmore, 2009), and others have implicated access to or the mapping of symbolic representations onto nonsymbolic ones (e.g., Noël & Rousselle, 2011; Rousselle & Noël, 2007). Several studies have shown that the effect of nonsymbolic magnitude representations is mediated by symbolic skills in both young children (vanMarle, Chu, Li, & Geary, 2014) and adults (Lyons & Beilock, 2011). A large recent meta-analysis (Chen & Li, 2014) shows an overall correlation of .20 for nonsymbolic magnitude measures with mathematics performance (in cross-sectional studies) and suggested that this result, coupled with power considerations, accounts for much of the empirical heterogeneity observed across individual studies with regard to the strength of relation between math and nonsymbolic magnitude, with some evidence of weaker relations (although still present) in older (>17 years) versus younger (<12 years) samples. Another recent meta-analysis (Fazio, Bailey, Thompson, & Siegler, 2014) shows a stronger relation of nonsymbolic magnitude to math performance in younger children ($r = .40$) than in older children ($r = .17$) or adults ($r = .21$).

Among domain-general predictors, the strongest evidence is for working memory (e.g., Alloway, Gathercole, Kirkwood, & Elliott, 2009; Gathercole, Brown, & Pickering, 2003; Geary, 2004, 2013; Raghubar et al., 2010). For instance, working memory is required for mathematics to keep in mind the multiple steps needed to complete a procedural computation. Other related processes important for mathematical skill include visuospatial functioning and language. For example, Szucs, Devine, Soltesz, Nobes, and Gabriel (2014) described a model in which visual–spatial working memory measures are key predictors in 9-year-olds' math performance (as opposed to number sense measures or pure spatial measures). Visual–spatial skills may also be important when math problems contain graphs or figures and for more complex math (e.g., geometry) where spatial skills are invoked more directly. An important role for language, often in the form of phonological awareness in young children, has also been implicated (e.g., Cirino, 2011; Koponen, Mononen, Rasanen, & Ahonen, 2006; Krajewski & Schneider, 2009; Martin, Cirino, Charp, & Barnes, 2014; Vukovic & Lesaux, 2013). Counting skills also rely in part on language (Koponen, Salmi, Eklund, & Aro, 2013; Koponen et al., 2007), and there is a predictable progression from counting to arithmetic (Geary, 2004).

Foundations for and prediction of later developing math skills

Relative to the volume of literature regarding math facts, whole number computation, and word problems, much less is known about fractions, proportional reasoning, and broader math skill at older ages. Therefore, the current study focused on later developing mathematics skills, when student mastery of math facts and basic whole number computation is expected, and therefore we expected these to comprise an additional layer of foundational skills. Because math develops in hierarchy, the conceptual relevance of procedural computational skills is clear. Facility with basic facts and multistep whole number operations aids all later mathematics because these are often needed at some point in order to solve fractions, proportions, and complex word problems.

The role for the understanding of more basic concepts of numerosity, and of domain-general cognitive skills, for later developing mathematics skills may at first seem unclear. However, both of these sets of skills have a conspicuous role, although for different reasons. The core ideas of numerosity are reflected in Siegler and Lortie-Forgues' (2014) integrative theory of numerical development. In this view, numerosity represents key milestones of numerical representation, beginning with magnitude representation, linking symbolic and nonsymbolic representations, and extending this understanding to include larger numbers, negative numbers, and the space between whole numbers (fractions and decimals). From this point of view, the stronger the individual and relative representations, the more easily those representations and their corresponding quantities might be evaluated and manipulated through mathematical operations. This understanding of quantity, thus, can act as a bridge between whole number and fraction performance (Siegler, Thompson, & Schneider, 2011). Because fractions and proportions involve ratios (a task that inherently requires a comparison of two magnitudes based on their ratio), representations may play a key role in operations that involve them. Geary, Hoard, Nugent, and Rouder (2015) provided evidence for a nuanced view in relation to algebraic cognition, with nonsymbolic comparison being more strongly related to numerical or spatial–numerical positioning, but less so to the proper structure of algebra equations or to more general algebraic achievement, at least when a strong set of covariates is included.

In Geary's (2004) framework, language and visuospatial systems support the representation of material, including (in the case of the latter) that of magnitude and the mental number line. The central executive controls (typically including working memory) attentional and inhibitory processes that are needed to perform procedurally complex computations. Further links are apparent in the model of von Aster and Shalev (2007), where increasing working memory resources are used in part to develop a strong mental number line, which itself serves as a representational redescription of a more general sense of magnitude.

In all, because of these underlying and overlapping contributions to our ability to represent, manipulate, and operate on magnitudes, it is feasible that domain-general skills (e.g., language, visuospatial skills, working memory), as well as individual differences in the development of numerosity (e.g., number line estimation, symbolic and nonsymbolic magnitude comparison), assist in the development of whole number skills, including fact retrieval and procedural computations, and may also contribute directly or indirectly through those whole number skills to support later developing mathematical skills, including fractions and proportional reasoning.

Fractions

For fraction performance in particular, understanding has been rapidly improving (Siegler, Fazio, Bailey, & Zhou, 2013). Recent studies have focused on the nature/types of fraction competencies (e.g., Hallett, Nunes, & Bryant, 2010; Hecht & Vagi, 2010; Hecht et al., 2007), on the prediction of fraction performance (Booth, Newton, & Twiss-Garrity, 2014; Hecht & Vagi, 2010; Jordan et al., 2013; Siegler et al., 2012; Vukovic et al., 2014), and on intervention (Fuchs et al., 2013; Gabriel et al., 2012). Prediction studies are most common, although their focus varies in terms of grade range and the kinds of predictor and fraction skills considered. A common focus for such studies is later elementary school, when fractions become a major component of the school curriculum.

Hecht and Vagi (2010) examined outcomes of fraction estimation, computation, and word problems in Grades 4 and 5; growth in these outcomes was predicted by whole number skills (arithmetic retrieval), domain-general skills, and other covariates (e.g., inattention, IQ, word reading) to varying degrees beyond fourth-grade math skills. Conceptual knowledge of fractions later predicted all fraction outcomes, but earlier procedural competence was weakly related to later conceptual competence. Jordan and colleagues (2013) and Vukovic and colleagues (2014) focused on a similar grade range. In Jordan and colleagues' study, Grade 3 inattention, number line estimation, and calculation fluency were predictive of Grade 4 conceptual and procedural fraction skills, whereas nonsymbolic comparison was not predictive of either outcome. Furthermore, working memory was uniquely related to procedural (but not conceptual) fraction outcomes, whereas language, reading fluency, and nonverbal reasoning were predictive of conceptual (but not procedural) fractions. In Vukovic and colleagues' study, some (but not all) domain-general skills were important for fraction concepts, albeit indirectly (via whole number computation or number line placement). Number knowledge (fluency in processing numerical magnitudes with both symbolic and nonsymbolic stimuli) showed a direct effect from Grade 1 to Grade 4. In that study, whole number *procedural* skills were also predictive of fraction *concepts* 2 years later.

Few studies have closely examined fraction performance beyond elementary school. An early exception was Fennema and Tartre (1985), who showed that spatial visualization was related to fractions in middle school students. More recently, Siegler and Pyke (2013) found in a small sample that differences between lower and higher achieving students in fractions at sixth grade grew wider by eighth grade. For fraction arithmetic, strong predictors were reading, whole number division, and fraction magnitude but not executive functioning (working memory and inhibition).

Across grades, these studies suggest that some combination of domain-specific and domain-general processes predicts procedural and conceptual fractions outcomes and that numerosity or whole number computation may mediate the effects of early domain-general processes on later fraction performance. Yet, inconsistency in findings demonstrates the need for more studies of this nature. This is especially the case given that fraction performance combines with whole number and conceptual knowledge in the prediction of algebra. Siegler and colleagues (2012) found in one large dataset that fractions, division, and multiplication in particular were strong predictors of algebra performance 5 years later. In the same study with a second large dataset, fraction and division performances were predictive of algebra (although all items were derived from the Woodcock–Johnson Applied Problems subtest). Other studies have shown the importance of domain-general skills for algebra outcomes, such as working memory (Lee, Ng, & Ng, 2009; Tolar, Lederberg, & Fletcher, 2009), although few studies simultaneously investigated all of the potential predictors, as was done in the current study.

Thus, the current study may provide bridging data to help link findings from the many studies of early developing math skills to the fewer studies of later developing math skills by focusing on a grade range in between early elementary school and high school and on outcomes in between whole number arithmetic and algebra. Because the larger study from which the current data were derived prohibited exhaustive evaluation given the setting, we elected to evaluate fraction performance with a broad measure that included both procedural and conceptual items but that focused primarily on computational problems rather than word problems (Brown & Quinn, 2007a, 2007b). We chose this measure given its demonstrated relation to algebra performance (and because the current study was completed just prior to the several recent studies referenced above).

Proportional reasoning

Proportional reasoning is potentially also important for algebra because both require the comparison of equivalency between two rational numbers. However, fewer studies have focused on proportional reasoning than on fractions. Gabel, Sherwood, and Enochs (1984) found that low proportional reasoning negatively impacted chemistry performance. More relevant, Kwon, Lawson, Chung, and Kim (2000) showed that measures of inhibiting ability, planning ability, figure/ground, and working memory capacity are predictive of proportional reasoning in adolescents. Recently, Kingston and Lyddy (2013) found self-efficacy and short-term memory to predict proportional reasoning. It is not clear whether predictors of individual differences in proportional reasoning are similar to or different from those of fractions or how either are related to general mathematics performance more generally following elementary school.

There are few well-researched measures of proportional reasoning available in the literature. However, Misailidou and Williams (2003) systematically developed a unidimensional instrument using or adapting an assortment of items previously used in curricular contexts and literature, each of which requires the determination of a missing value from a proportion given three other values. They developed a total of 38 items (25 with a graphical model and 13 without) and administered short overlapping forms to a large group of students aged 10 to 14 years. From these, they developed two parallel 13-item diagnostic measures that fit the original Rasch structure. Given the strong development of this measure, it was adopted for the measure of proportional reasoning used here.

Broad mathematics

Of the studies that have examined general mathematical performance, few systematically include a range of both specific processes (e.g., numerosity, whole number arithmetic) and general processes (e.g., working memory, spatial skills, language), although studies do address these separately. For example, Fazio and colleagues (2014) found a strong influence for (particularly symbolic) magnitude knowledge (of whole numbers or fractions) in predicting general math achievement (a state test) in fifth graders but did not include domain-general predictors. Although not a predictive study, Matteson (2006) found that algebraic problems on a state test across Grades 3 to 8 were most often represented verbally (or to a lesser extent graphically) as opposed to numerically or symbolically, although the reverse was true of the required solutions. Only two studies included a variety of types of predictive measures. Bailey, Hoard, Nugent, and Geary (2012) found that fraction comparison in Grade 6 predicts Grade 7 math, but the reverse was not true. With all factors considered, the strongest predictors of Grade 7 math did not include fraction comparison but rather computational fractions and computational arithmetic as well as number (number line and fluency in processing numerical magnitudes with both symbolic and nonsymbolic stimuli) and IQ. In Siegler and Pyke (2013), for general mathematics (state test), strong predictors included reading, executive functions, whole number division, and fraction magnitude.

The current study

The current study sought to clarify the relative roles of domain-specific and domain-general contributions to middle school mathematics given that both have crucial importance for later developing skills including algebra. The current study contributes to the literature in four ways. First, we extend knowledge regarding the relative roles of domain-specific (numerosity) and domain-general (cognitive) contributors to math skills to an understudied age range. Second, within the numerosity domain, symbolic and nonsymbolic predictors are distinguished; within the cognitive/neuropsychological domain, language and visuospatial predictors are each considered. Third, three different types of math skills relevant in middle school are considered: fraction performance, proportional reasoning, and a state mathematics assessment. Finally, and perhaps most important, a distal to proximal hierarchy is evaluated, where the effects of these predictors on outcomes are evaluated both directly and indirectly via numerosity and/or elementary arithmetic skills.

We hypothesized that both domain-specific and domain-general processes would be relevant across outcomes. This is consistent with studies demonstrating that although domain-general skills are involved in basic facts, their role grows in importance as the complexity of the mathematics outcome skill increases (e.g., from computations to word problems; Fuchs et al., 2010a, 2010b). The literature supports the role of numerosity for basic skills, although less information is available regarding this role in middle school mathematics performance.

We go beyond this basic hypothesis to more specifically reflect the hierarchy of math development. For fraction competency, we expected that the mathematics skills most proximal to the outcome (whole number math facts and procedural computation) would exhibit the strongest direct effects. We hypothesized that numerosity (e.g., magnitude/number line estimation, symbolic and nonsymbolic comparison) would have both direct and indirect effects. We expected domain-general skills (e.g., language, working memory, spatial) to also have both direct and indirect effects, particularly working memory. Limited information is available regarding proportional reasoning, but given similar complexity for fractions and proportional reasoning (and because proportional reasoning is involved in fraction performance; Vukovic et al., 2014), we expected results to be similar to those for fractions. For broad mathematics (state test), we again expected similar findings, but with the inclusion of fractions and proportional reasoning as predictors, we expected the role of more distal measures used in other models to be reduced. This could manifest by finding only indirect effects for numerosity and domain-general skills in the context of whole number computation, fractions, and proportional reasoning.

Method

Participants

Participants were 162 sixth-grade students, from three schools in a large southwestern metropolitan area of the United States, whose average age was 12.4 years ($SD = 0.53$). Slightly more than half

(53%) of the participants were male, and most students were Hispanic/Latino (75%) or African American (20%). All students were instructed in English, although 16% were classified as limited English. Most students were in regular classes, with 1% classified as gifted/talented and 7% receiving special educational services. Reduced/free lunch status was not available from one school, but proportions of economic disadvantage at all locations were high (~80% at the school level and 79–88% for individual students at the schools where data were available). These demographic characteristics were consistent with the individual schools and districts as a whole. Several of these descriptive variables were considered as covariates in analyses (see below). Written informed consent was obtained from parents, as was assent from students.

Measures

Table 1 shows means, standard deviations, skew, and kurtosis for measures in their original metrics.

Domain-general (cognitive/neuropsychological) measures

Spatial working memory was assessed with the Automated Symmetry Span (Kane et al., 2004; Unsworth, Redick, Heitz, Broadway, & Engle, 2009), which requires students to maintain a series of spatial locations (recall task) in memory while making veridical judgments about the symmetry of figures (processing task). Participants are instructed to ascertain symmetry as quickly and accurately as possible and to maintain a "correct symmetry" score of at least 85%. There are 12 trials, with 3 trials each at span lengths of 2, 3, 4, and 5; the order of each trial is randomly generated. The primary measure used is a "partial credit load" (PCL) score, which is more reliable than all-or-nothing scoring (Conway et al., 2005). In PCL scoring, points are assigned for each trial according to the span length, so the maximum possible score is 42. Four individuals did not have data on the last of the 12 trials due to computer administration error, and their scores on this trial were prorated. In this sample, α was .79.

Language was assessed with the Wechsler Abbreviated Scales of Intelligence (WASI) Vocabulary subtest (Wechsler, 1999), which requires students to define orally presented words and, as such, is a measure of expressive vocabulary. Reliability is excellent for this subtest (α = .86–.90; Wechsler, 1999). Students also received the Visual–Auditory Learning task of the Woodcock–Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001). This task requires students to associate a rebus to a word and "read" sentences using the symbols. This task assesses memory but in the context of symbolic association and syntactic manipulation. It requires approximately 10 min, and reliability among 11- to 15-year-olds is good (range of α = .76–.90; McGrew & Woodcock, 2001). Age-based standard scores were used as the primary variable from each measure. Both are well standardized and commonly used in clinical settings. The two measures correlated r = .48 and were formed into a language composite for parsimony and given similar expectations for their relations across the outcomes considered here. Three individuals were outliers in terms of their composite, and these scores were trimmed to above the next highest value to maintain their rank ordering (and correlations with outcomes were essentially unchanged relative to original values).

Visuospatial rotation was assessed with the Revised Vandenberg & Kuse Visuospatial Rotation Test (MRT; Peters, 1995), which is a redrawn version of the Vandenberg and Kuse (1978) figures. Each of the items has a target figure and four related stimuli; participants must identify which two of the four stimuli are rotated versions of the target. Participants worked with up to 24 items for 6 min. This measure loaded highly on a spatial visualization latent factor (.83) and was correlated with measures of algebra procedural skill (.28–.30) among college students (Tolar et al., 2009). Reliability in this sample was α = .76. The score used was the number of items answered correctly; the maximum was 24 because to complete an item correctly, participants needed to identify both of the rotated versions of the stimulus figure. Three individuals were outliers, and these scores were trimmed to above the next highest value to maintain their rank ordering (and correlations with outcomes were essentially unchanged relative to original values).

Table 1
Descriptive statistics of (original) measures (N = 150–162).

Area	Measure	Scale	Mean (*SD*)	Skew	Kurtosis
Domain general					
	Working memory	Raw (0–42)	21.19 (8.45)	−0.48	−0.04
	WASI vocabulary	Scale score (50/10)	44.97 (8.97)	−0.49	0.46
	WJ-III visual auditory learning	Scale score (100/15)	90.29 (12.68)	−0.56	1.20
	Visuospatial rotation	Raw (0–24)	4.5 (3.09)	1.26	2.81
Domain specific (number)					
	Nonsymbolic comparison	Weber fraction	0.24 (0.10)	3.05	13.93
	Symbolic comparison	Mean correct response time (ms)	798 (249)	2.24	7.36
	Number line estimation	Mean absolute deviation (cm)	81.76 (51.28)	2.56	8.17
Fact-based retrieval					
	Single digit addition	Raw (0–99)	28.27 (8.70)	1.42	4.32
	Single digit multiplication	Raw (0–99)	31.62 (13.89)	0.08	0.66
Procedural computation					
	ETS subtraction and multiplication	Raw (0–60)	12.04 (7.03)	−0.06	0.37
	ETS addition and subtraction	Raw (0–60)	17.58 (5.66)	0.35	0.90
Grade 6 mathematical outcomes					
	Fraction competency	Raw (0–25)	4.93 (2.92)	0.80	0.43
	Proportional reasoning	Raw (0–10)	3.06 (1.98)	0.55	0.08
	State test	Raw (0–52)	32.00 (10.55)	−0.13	−0.90

Note. The data show original values. Transformation/outliers (typically minor and/or few) significantly improved distributions of visuospatial rotation, nonsymbolic comparison, symbolic comparison, and number line estimation. See text for descriptions. The combined variables of language (vocabulary and visual auditory learning), fact-based retrieval, and procedural computation all showed good distributional properties. WJ-III, Woodcock–Johnson II I; ETS, Educational Testing Service.

Domain-specific (numerosity) measures

Numerosity was assessed with three measures. Panamath is a research-based, downloadable computerized measure (Halberda et al., 2008; http://www.panamath.com) and is widely used as a measure of acuity of the approximate number system (Hyde, Khanum, & Spelke, 2014; Jordan et al., 2013; Libertus, Odic, & Halberda, 2012). Students see dots on a computer screen (yellow on the left and blue on the right) for 600 ms, followed by a backward mask and then a fixation symbol; students decide as quickly as possible which side has more dots by pressing a key on the left or right side of the keyboard. Dot size, area, and the ratio between the dot set sizes vary systematically. For the purposes of the current study, 200 trials were administered and varied the extent to which dot size and area were consistent versus inconsistent with numerosity as well as the ratio of dots to one another. The key measures are percentage correct, response time (RT), and discriminability (Weber fraction, where lower fractions indicate the ability to discriminate dot sets that appear similar). Four individuals were outliers in terms of their Weber fraction, and their scores were trimmed to above the next highest value to maintain their rank ordering (and correlations with outcomes were essentially unchanged relative to original values). Across the different types of trials, Weber fraction scores generated from each correlated >.80 with one another and >.90 with their corresponding percentage correct scores.

Symbolic Comparison has 91 trials in which students see two Arabic numerals (either single digit or double digit). The task is to decide which is greater as quickly and accurately as possible. The ratio of the two numbers was designed to maximally overlap with those of the nonsymbolic presentation of dots (i.e., Panamath). The primary variables are accuracy and RT (mean, median, and standard deviation across trials), although as expected accuracy was high (mean of 85/91 correct) except for one student, who was excluded for this measure due to chance performance. In this sample, α = .98 for RT. For the three indexes of RT, three individuals were outliers (slower than others); these data were trimmed to above the next highest value to maintain their rank ordering. These scores were also log transformed due to some skewness. The three indexes were strongly correlated and similarly distributed; therefore, only mean RT (trimmed and transformed) was used.

The Number Line Estimation task (Booth & Siegler, 2006; Siegler & Booth, 2004) is strongly predictive of math achievement. Students see an Arabic numeral and a line anchored with the numerals 0 and 1000, and they place the numeral on the number line. The score was the mean deviation from the number's actual line position (Booth & Siegler, 2006; Siegler & Booth, 2004; Siegler & Opfer, 2003). Raw scores were positively skewed and kurtotic (a significant minority of students had larger error scores). Therefore, a log-transformed measure was used in analyses. Correlations with outcome measures did not differ whether the variable was log transformed or whether participants with the most deviant scores were trimmed or deleted. Reliability in this sample was $\alpha = .88$.

Whole number retrieval and computation

Two measures of *fact-based retrieval* were used, each with a similar format. Problems appeared across two pages with 17 rows and 6 items per row for a total of 99 problems for each (the last row had 3 items), and participants had 1 min to complete as many problems as possible. For Single Digit Addition, problems comprised all possible combinations of the addends 0 through 9; Single Digit Multiplication multiplicands and multipliers ranged from 0 to 9. For both measures, problems were organized systematically and then randomized for presentation. The score was the number of correctly completed problems within the time limit minus the number of incorrect responses (although errors were relatively uncommon). The two measures were combined to form a math fact retrieval composite; the correlation between them was $r = .53$.

Two measures from the ETS Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, Harman, & Derman, 1976) were used to assess *procedural computation.* The Subtraction and Multiplication Test is part of the Number Facility factor and consists of alternating rows of subtraction and multiplication problems. Minuends and subtrahends in all subtraction problems are two digits (e.g., $98 - 75$). Multiplication problems have two-digit multiplicands and single-digit multipliers (e.g., 86×6). This test has two parts, but only the first part, with 60 items, was completed with a 2-min time limit. Reliability is .92 for sixth-grade students and ninth-grade boys (Ekstrom et al., 1976). The Addition and Subtraction Correction Test consists of 60 items in two rows, showing two-digit numbers and a potential sum in horizontal format (e.g., $26 - 14 = 13$); students identify whether the problem and its sum is correct. The score for each measure is the number of correct items minus the number of incorrect responses. The two measures were combined to form a whole number procedural computation composite; the correlation between them was $r = .67$.

Mathematical outcomes

Three additional measures assessed *fraction performance, proportional reasoning,* and *general mathematics.* The Brown and Quinn Fraction Competency Test (Brown & Quinn, 2007a, 2007b) is a 25-item broad measure of fractions meant to assess conceptual knowledge and procedures, with presentation primarily computational and only a few word problems (e.g., "Find the sum of 5/12 and 3/8"; "If n gets very large, then $1/n$ gets ... [very close to 1, very close to 0, very large too]"). The full items are provided in Brown and Quinn (2007a, 2007b) and were rearranged here by the order of difficulty found in the original study. Items are classed into six categories: algorithmic applications, application of basic fraction concepts in word problems, elementary algebraic concepts, specific arithmetic skills that are prerequisite to algebra, comprehension of the structure of rational numbers, and computational fluency. However, the authors noted that that items could fall into more than one category (Brown & Quinn, 2006); in addition, the number of items in each group was rather small and the groups of items were not described as subtests, nor were separate reliability statistics provided. Among 138 first-year algebra students (predominantly Grade 9), total performance accuracy was 52%, with $r = .58$ with their algebra course final exam (Brown & Quinn, 2007a, 2007b). Students were given 10 min to complete as many items as possible. Reliability in this sample was $\alpha = .69$. The total score was used for analyses.

The Diagnostic Assessment of Proportional Reasoning (Misailidou & Williams, 2003) has 13 items calibrated with 303 students (ages 10–14 years) using item response theory (IRT) techniques. Correlation between item difficulty estimates was $r = .93$ between the two samples. Item selection, development, and other psychometric properties are described in some detail in the original publication, and all items are presented explicitly (Misailidou & Williams, 2003). From this item pool, 13 items were scaled to a diagnostic test that could serve as an assessment of ratio attainment as well as to pull for additive errors in solving ratio and proportion problems. This 13-item test was cross-validated on a second sample ($N = 212$, ages 10–13 years); for the current study, 10 of these 13 content-unique items were used, with thresholds (difficulty parameters in the IRT context) that ranged from −2.72 to +2.01. The following is an example item ($b = .65$): "10 campers camped at the 'Blue Mountain' camp the previous week. Each day the camp's cook baked eight loaves of bread for them to share. This Monday 15 campers will camp at the 'Blue Mountain' camp. How many loaves will the cook need to bake each day?" A total score is available for analyses, consistent with its development. Students were given 10 min to complete as many items as possible. Because the assessment was developed in the United Kingdom, unfamiliar language was modified (e.g., for a money question, changing "pounds" to "dollars"). Reliability in this sample was $\alpha = .69$.

Finally, the state standardized assessment for mathematics in Grade 6 was used. At this grade, there are 52 questions across five sections: (a) numbers, operations, and quantitative reasoning (16 items); (b) patterns, relationships, and algebraic reasoning (12 items); (c) geometry and spatial reasoning (8 items); (d) measurement (8 items); and (e) probability and statistics (8 items). The total raw score was used. The five individual portions of this assessment correlated in the range of $r = .53$ to .73.

Procedures

Students were assessed in their schools during late spring of 2012 at times convenient to their schedules and to school staff/personnel. Total assessment time was approximately 2.5 h, including breaks as needed. Test order was quasi-randomized, primarily for logistic reasons, including the length of the battery, the computer equipment needed for some (not all) measures, and the school setting. This prevented both a fixed order and a regimented random ordering. Instead, examiners (monitored by on-site supervisors) were encouraged to begin at different points and to work in forward or reverse order, depending on whether computers were involved. Examiners were trained and supervised by the authors, and most examiners were experienced both with children in schools and with standardized test administration and similar measures on other projects. In addition, examiners participated in a full-day training session that included didactic presentations and practice. Examiners then practiced independently and with other research staff members and passed a "checkout" procedure with the testing coordinator prior to actual testing. On-site supervision was also provided.

Analysis plan

Preliminary analyses included examination of variable distributions, outliers, and influential observations within and across conditions described above and reported in Table 1. Next, covariates of age, sex, ethnicity, language status (limited English), and educational status (receiving special education services or not) were considered. Age was significantly (negatively) related to each outcome. Sex was unrelated to primary outcomes and so was not further considered. Ethnicity (Hispanic/Latino or not [predominantly African American]) was related to each outcome, with students who were Hispanic/Latino outperforming those who were not. Language status was correlated with fractions ($p < .015$) and proportional reasoning ($p < .046$) outcomes, where it accounted for a small proportion of variance; language status was unrelated to the state test. Similarly, educational status (e.g., special education) was related to proportional reasoning ($p < .006$) and to the state test ($p < .010$) but not to fractions. When age, ethnicity, special education, and language status were considered together, age and ethnicity (and language status for fractions and proportional reasoning) were significant predictors of outcomes, accounting for 11% to 21% of the variance in outcomes. Finally, models were run with and without these covariates, with unique contributions for age and ethnicity (but not language status) for proportional reasoning and state test outcomes; the addition of these covariates did not alter (enhance or ameliorate) the contributions of the cognitive, number, or arithmetic variables to the models and did not interact with them. Therefore final models are presented without covariates.

The primary analytical approach was path analyses via Mplus (Muthen & Muthen, 2007). Although we used only observed variables (and composites), we chose this approach over multiple regression (or hierarchical regression) even though they are related. We did so because we could simultaneously estimate direct and all manners of indirect effects, and we could test the model (which has degrees of freedom because some paths are constrained to zero) as a whole. Doing so more accurately represents our hypothesis of a hierarchy of math skill and its influences. In this way, direct effects of variables offer an index of a variable-unique contribution, considering all other variables (similar to what would be obtained in a hierarchical regression), but also evaluates the extent to which cognitive and numerosity variables operate through whole number arithmetic. In path analysis with some degrees of freedom, there are fit indexes at both the global level (e.g., chi-square [χ^2], comparative fit index [CFI], Tucker–Lewis index [TLI], root mean square error of approximation [RMSEA], standardized root mean square residual [SRMR]) and the local level (parameter estimates). Furthermore, if data are missing for endogenous variables (which was generally rare in this study, with covariance coverage typically being >95%), model-implied covariances are based on all available participants and global fit is assessed via maximum likelihood; only individuals missing data on exogenous variables ($n = 3$), in our case the cognitive variables, were excluded. In addition, direct and indirect effects can be produced all at once rather than computed by hand. Finally, bootstrapped confidence intervals are computed (because indirect or mediational effects are often non-normally distributed; Preacher & Hayes, 2004, 2008; Shrout & Bolger, 2002), and this approach can extend more traditional perspectives (Hayes, 2009; MacKinnon & Fairchild, 2009). The path analytic framework is important for this study because several of the proposed indirect effects operate through more than one intervening variable. Further details regarding path analyses, mediation, and related techniques may be found in relevant texts and reviews (e.g., Klein, 2010; MacKinnon, 2008).

The interpreted model for each outcome evaluated the overall schema hypothesized here, in line with the hierarchy of math development. Here, cognitive variables (language, working memory, and visuospatial rotation) support numerosity variables (number line, symbolic comparison, and nonsymbolic comparison); then, cognitive and numerosity variables support fact retrieval and procedural computation, and each variable predicts math outcomes (fractions, proportional reasoning, and general math achievement), either directly or indirectly via more proximal effects. Indexes of both global fit (overall model) and local fit (variable parameter estimates) were evaluated. Regarding the latter, we specifically evaluated total, direct, total indirect, and specific indirect effects, in line with our hypothesized variable ordering. In the text, and in the tables and figures, we present effects as standardized regression coefficients (β) to help ease interpretability. Direct effects are analogous to unique effects in a standard multiple regression analysis. Specific indirect effects are the result of multiplying the coefficients along a given route to the outcome (e.g., the coefficients of the paths from language to number line times the coefficient of the path from number line to fact retrieval and so on to procedural computation and to fractions); the most "distal" variables (i.e., the cognitive variables) have the highest number of specific indirect effects. The cumulative or total indirect effect is the sum of all the specific indirect effects. The total effect is the sum of the cumulative indirect and direct effects.

Results

Correlations among variables are found in Table 2. These show moderate to strong relationships among outcomes, with similarly strong relationships between outcomes and the whole number variables, in particular procedural computation.

Fractions

The first descriptive step simply included all predictors (as noted, results were highly similar when covariates were considered) and so is a multiple regression. The predictors accounted for $R^2 = 47\%$ of the variance in fraction performance, with unique contributions for whole number (procedural computation), numerosity (symbolic comparison), and visuospatial rotation. The first overidentified model (with domain-general variables predicting numerosity, both of these predicting whole number computation/retrieval, and all variables predicting fractions performance) was ill-fitting (e.g., RMSEA = 0.371, CFI = 0.751); however, this was due to a strong, unaccounted-for relation between the intermediary whole number variables (math facts and procedural computation). Given that math

Table 2
Correlations of variables used in path models (N = 150–162).

	1	2	3	4	5	6	7	8	9	10	11
1. Working memory	1.00										
2. Language	.226	1.00									
3. Visuospatial rotation	.194	.121	1.00								
4. Nonsymbolic comparison	−.136	−.044	−.005	1.00							
5. Symbolic comparison	−.252	−.302	−.139	.078	1.00						
6. Number line estimation	−.164	−.214	−.122	.162	.133	1.00					
7. Fact-based retrieval	.276	.193	.060	−.139	−.495	−.345	1.00				
8. Procedural computation	.325	.311	.142	−.056	−.606	−.301	.769	1.00			
9. Fraction competency	.253	.288	.201	.048	−.307	−.264	.466	.650	1.00		
10. Proportional reasoning	.296	.355	.176	−.155	−.366	−.318	.474	.551	.498	1.00	
11. State test	.265	.372	.247	−.248	−.323	−.418	.438	.542	.421	.681	1.00

Note. Correlations with absolute values > .160 are $p < .05$; correlations > .200 are $p < .01$; correlations > .263 are $p < .001$.

facts are clearly developed prior to procedural computations in the math hierarchy, we related these by including a step (from math facts to procedural computations); indirect effects for this additional step were also considered. This model fit very well (χ^2 = 3.928, df = 3, CFI = 0.997, RMSEA = 0.044, 90% confidence interval [CI] = 0.000–0.148; SRMR = 0.025), R^2 = 47%, but revealed no total, direct, or indirect effects involving visuospatial rotation or nonsymbolic comparison.

Significant direct paths for this model appear in Fig. 1 but were apparent only for procedural computations; select (significant and cumulative) indirect path statistics appear in Table 3 (under Model 1). The impact of specific predictors is read as follows. Because visuospatial rotation and nonsymbolic comparison showed no direct or indirect effects, there is no line from these variables to fractions in Fig. 1, and only the total indirect effects, which are not significant (with the total number of indirect paths for each variable listed in parentheses), appear in Table 3 (in italics). Language showed a different combination of effects, although the direct path is not displayed in Fig. 1 because, as noted, this effect was not significant. There was a strong total effect (β = .226) and cumulative indirect effect (β = .141). The two significant specific indirect paths appear in Table 3, each of which operated through procedural computations (in addition to others). Spatial working memory showed a similar pattern, that is, no direct effect but strong total (β = .175) and cumulative indirect (β = .139) effects; the only significant specific indirect effect was the route through symbolic comparison, fact retrieval, and procedural computations. Symbolic comparison also showed no direct effect but significant total (β = −.202) and total indirect (β = −.343) effects; two significant specific indirect paths operated through procedural computations. Number line showed no direct effect either, although with significant total (β = −.184) and cumulative indirect (β = −.126) effects; one significant specific indirect path via math fact retrieval and procedural computations appears in Table 3. Finally, there was no direct effect of fact retrieval, but the total effect (β = .371) and cumulative indirect effect (there are no separate specific indirect effects because only one is available) were significant.

Proportional reasoning

Again, the first descriptive multiple regression step simply included all predictors (with similar results with and without covariates considered), which accounted for R^2 = 40% of the variance, with unique contributions only for whole number (procedural computation), language, and numerosity (number line). Fact retrieval was again added as a path between numerosity and procedural computations. The resultant model fit very well (χ^2 = 3.579, df = 3, CFI = 0.998, RMSEA = 0.035, 90% CI = 0.000 to 0.143, SRMR = 0.024), R^2 = 38%, but similar to the fractions model revealed no total, direct, or indirect effects involving visuospatial rotation or nonsymbolic comparison.

Significant direct paths for this model (language and procedural computations) are displayed in Fig. 2; significant specific indirect (and all cumulative indirect) path statistics appear in Table 3 (under Model 2). Again, all indirect paths operated through procedural computations, and in turn all paths to

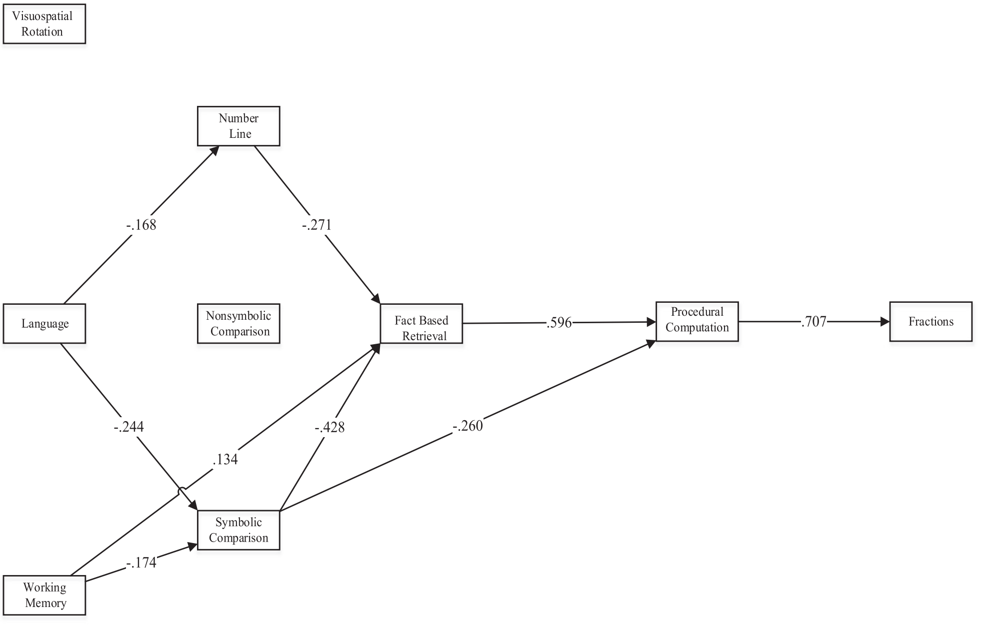

Fig. 1. Path model predicting fraction competency. Values shown are standardized beta weights. Only significant paths are shown for clarity. Standard errors and residuals are not shown for space reasons. Select indirect path statistics (all significant indirect paths and all cumulative indirect effects) appear in Table 3. All specific significant indirect effects in Table 3 are computable from the values in the figure.

procedural computations operated through math facts or symbolic comparison. As noted, visuospatial rotation and nonsymbolic comparison showed no direct or indirect effects. Language showed a strong direct effect ($\beta = .183$; see Fig. 2) and total effect ($\beta = .303$). The cumulative indirect effect was significant, with two significant specific indirect paths (see Table 3). Considering spatial working memory, number line, and symbolic comparison, none showed significant direct effects, but each showed total effects ($\beta = .209$, $\beta = -.205$, and $\beta = -.221$, respectively) and cumulative indirect effects ($\beta = .138$, $\beta = -.095$, and $\beta = -.222$, respectively). As indicated in Table 3, these variables had zero, one, and two significant specific indirect effects, respectively. For fact retrieval, similar to fraction performance, there was no direct effect, but the total ($\beta = .306$) and cumulative indirect effects were significant.

General mathematics

The descriptive model accounted for $R^2 = 46\%$ of the variance, with unique contributions for visuospatial rotation, language, numerosity (number line estimation and nonsymbolic comparison), and procedural computations. The overall model (considering direct and indirect effects, including those from math facts and procedural computation, as per the models above) fit extremely well ($\chi^2 = 3.263$, $df = 3$, CFI = 0.999, RMSEA = 0.023, 90% CI = 0.000 to 0.138; SRMR = 0.024), $R^2 = 44\%$.

Significant direct paths for this model were the same as those found for the descriptive regression model and are displayed in Fig. 3; significant specific and cumulative indirect path statistics appear in Table 3 (under Model 3). The *total* effects for each variable in the model were significant (spatial working memory, $\beta = .167$; language, $\beta = .319$; visuospatial rotation, $\beta = .183$; number line, $\beta = -.294$; symbolic comparison, $\beta = -.158$; nonsymbolic comparison, $\beta = -.172$; fact retrieval, $\beta = .239$). Cumulative indirect effects were significant for each predictor except visuospatial rotation and nonsymbolic rotation, and significant specific indirect effects were noted for language (three), number line (one), symbolic comparison (two), and fact retrieval (one) (see Table 3). Each of these specific indirect effects

Table 3
Standardized total indirect and significant indirect results.

Predictor	Indirect path	Estimate	Confidence interval
Model 1 (fraction performance)			
Visuospatial rotation	*Cumulative Indirect (15)*	*0.044*	*−0.051 to 0.139*
Language	Cumulative Indirect (15)	0.141	0.042 to 0.240
	→ Symbol Compare → Procedures →	0.045	0.007 to 0.083
	→ Symbol Compare → Facts → Procedures →	0.044	0.011 to 0.077
Working memory	Cumulative Indirect (15)	0.139	0.040 to 0.238
	→ Symbol Compare → Facts → Procedures	0.031	0.000 to 0.063
Nonsymbolic comparison	*Cumulative Indirect (3)*	*0.027*	*−0.073 to 0.126*
Symbolic comparison	Cumulative Indirect (3)	−0.343	−0.449 to −0.237
	→ Procedures →	−0.184	−0.282 to −0.085
	→ Facts → Procedures →	−0.180	−0.265 to −0.095
Number line	Cumulative Indirect (3)	−0.126	−0.216 to −0.035
	→ Facts → Procedures →	−0.114	−0.183 to −0.045
Facts	Cumulative Indirect (1, via Procedures)	0.421	0.272 to 0.570
Model 2 (proportional reasoning)			
Visuospatial rotation	*Cumulative Indirect (15)*	*0.031*	*−0.040 to 0.101*
Language	Cumulative Indirect (15)	0.120	0.035 to 0.204
	→ Symbol Compare → Procedures →	0.022	0.001 to 0.043
	→ Symbol Compare → Facts → Procedures →	0.022	0.003 to 0.040
Working memory	Cumulative Indirect (15)	0.138	0.059 to 0.217
Nonsymbolic comparison	*Cumulative Indirect (3)*	*0.009*	*−0.054 to 0.072*
Symbolic comparison	Cumulative Indirect (3)	−0.222	−0.307 to −0.138
	→ Procedures →	−0.091	−0.154 to −0.028
	→ Facts → Procedures →	−0.090	−0.149 to −0.030
Number line	Cumulative Indirect (3)	−0.095	−0.157 to −0.033
	→ Facts → Procedures →	−0.056	−0.102 to −0.010
Facts	Cumulative Indirect (1, via Procedures)	0.209	0.082 to 0.336
Model 3 (state test outcome)			
Visuospatial rotation	*Cumulative Indirect (15)*	*0.035*	*−0.043 to 0.112*
Language	Cumulative Indirect (15)	0.128	0.032 to 0.225
	→ Number Line →	0.037	0.002 to 0.071
	→ Symbol Compare → Procedures →	0.027	0.002 to 0.052
	→ Symbol Compare → Facts → Procedures →	0.026	0.004 to 0.049
Working memory	Cumulative Indirect (15)	0.148	0.057 to 0.238
Nonsymbolic comparison	*Cumulative Indirect (3)*	*0.017*	*−0.047 to 0.082*
Symbolic comparison	Cumulative Indirect (3)	−0.213	−0.302 to −0.124
	→ Procedures →	−0.110	−0.183 to −0.038
	→ Facts → Procedures →	−0.109	−0.180 to −0.037
Number line	Cumulative Indirect (3)	−0.079	−0.143 to −0.015
	→ Facts → Procedures →	−0.067	−0.120 to −0.014
Facts	Cumulative Indirect (1, via Procedures)	0.253	0.107 to 0.399
Model 4 (state test outcome with fractions and proportional reasoning)			
Visuospatial rotation	*n/a*	*n/a*	*n/a*
Language	Cumulative Indirect (14)	0.190	0.098 to 0.283
	→ Proportions →	0.099	0.033 to 0.165
	→ Number Line →	0.032	0.000 to 0.064
	→ Symbol Compare → Procedures → Proportions →	0.018	0.001 to 0.035
	→ Symbol Compare → Facts → Procedures → Proportions →	0.015	0.002 to 0.027
Working memory	Cumulative Indirect (12)	0.075	0.020 to 0.129
	→ Facts → Procedures → Proportions	0.019	0.000 to 0.039
	→ Symbol Compare → Procedures → Proportions →	0.014	0.000 to 0.029
	→ Symbol Compare → Facts → Procedures → Proportions →	0.012	0.001 to 0.022

Table 3 (*continued*)

Predictor	Indirect path	Estimate	Confidence interval
Nonsymbolic comparison	*n/a*	*n/a*	*n/a*
Symbolic comparison	Cumulative Indirect (7)	−0.288	−0.384 to −0.192
	→ Procedures →	−0.109	−0.195 to −0.023
	→ Procedures → Proportions →	−0.075	−0.114 to −0.035
	→ Facts → Procedures →	−0.088	−0.157 to −0.019
	→ Facts → Procedures → Proportions →	−0.060	−0.089 to −0.031
Number line	Cumulative Indirect (4)	−0.059	−0.105 to −0.013
	→ Facts → Procedures →	−0.046	−0.089 to −0.004
	Facts → Procedures → Proportions →	−0.032	−0.054 to −0.009
Facts	Cumulative Indirect (3)	0.329	0.191 to 0.467
	→ Procedures →	0.204	0.058 to 0.351
	→ Procedures → Proportions →	0.139	0.079 to 0.199
Procedures	Cumulative Indirect (2)	0.207	0.091 to 0.323
	→ Proportions →	0.232	0.142 to 0.322

Note. Nonsymbolic Comparison: Panamath (comparison of dot arrays of varying ratios, often used as a measure of approximate number system acuity); Symbolic Comparison: comparison of Arabic numerals of varying ratios; Number Line: estimation of location of a given Arabic numeral along an analog scale bounded by 0 and 1000; Facts: single-digit addition and multiplication; Procedures: ETS Subtraction/Multiplication and Addition/Subtraction Correction. Cumulative Indirect effects are shown for each predictor (summed over all possible indirect effects, noted in parentheses); those in italics are not significant (bootstrapped confidence intervals cross zero). Specific indirect effects that are not significant are not listed in the table (and, therefore, the cumulative indirect effect is not the sum of the specific indirect effects shown in the table). The number of specific indirect effects in Model 4 is fewer than that in other models because it was built from them. See text for description of pattern of direct effects and total effects; see figures for graphical display of direct effects.

operated through procedural computations except for the pathway from language through number line to state test performance.

The above analyses for general mathematics performance were repeated but including fraction performance and proportional reasoning in the model as additional predictors (see Fig. 4 and Table 3 under Model 4). The regression model showed unique effects for visuospatial rotation, numerosity (nonsymbolic comparison and number line estimation), procedural computations, and proportional reasoning (although not for language), accounting for R^2 = 58%. Because previous models had already evaluated the impact of cognitive, numerosity, and whole number variables on fractions and proportional reasoning, only those effects found in those previous models were included as follows. General mathematics was regressed on all predictors/covariates. Fraction performance was regressed only on procedural computations, and proportional reasoning was regressed on language and procedural computations for analogous reasons. Procedural computation was regressed on math fact retrieval and symbolic comparison, whereas math fact retrieval was regressed on symbolic comparison, number line estimation, and spatial working memory. Finally, symbolic comparison was regressed on language and spatial working memory, and number line was regressed only on language. Thus, all indirect effects operated via these routes; only direct effects were allowed for visuospatial rotation and nonsymbolic comparison given results of other models. This model fit was marginal (χ^2 = 53.090, df = 28, CFI = 0.952, RMSEA = 0.076, 90% CI = 0.044 to 0.107, SRMR = 0.074), R^2 = 54%, although it should be noted that this model was *substantially* more restrictive than any of the other models evaluated; direct effects again were the same as those of the regression analyses.

In this model, total effects were significant for all variables in the model except spatial working memory (language, β = .308; visuospatial rotation, β = .109; number line, β = −.242; symbolic comparison, β = −.202; nonsymbolic comparison, β = −.144; fact retrieval, β = .260; and procedural computations, β = .547). Cumulative indirect effects were present for all variables for which they were evaluated. Finally, significant specific indirect effects were noted for language (four), spatial working memory (three), number line (two), symbolic comparison (four), fact retrieval (two), and procedural computations (one), each of which appears in Table 3.

To summarize, in the model for the state test without fractions or proportional reasoning (see Fig. 3), all indirect effects operated through number line estimation, symbolic comparison, fact

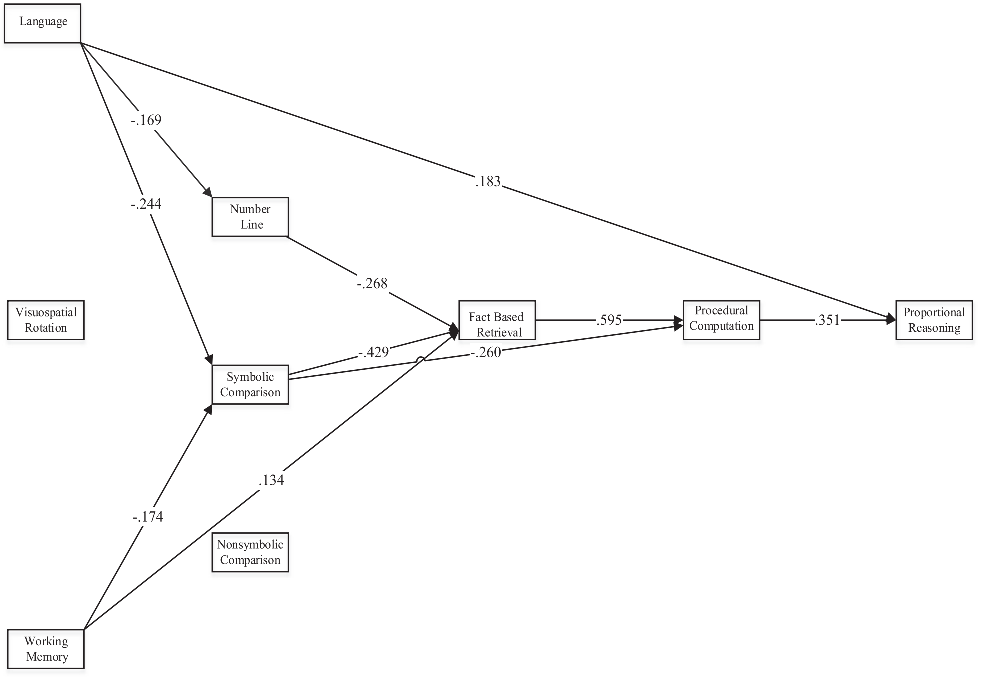

Fig. 2. Path model predicting proportional reasoning skill. Values shown are standardized beta weights. Only significant paths are shown for clarity. Standard errors and residuals are not shown for space reasons. Select indirect path statistics (all significant indirect paths and all cumulative indirect effects) appear in Table 3. All specific significant indirect effects in Table 3 are computable from the values in the figure.

retrieval, and/or procedural computations. When fractions and proportional reasoning were added, direct effects were highly similar (except that in this model proportional reasoning was a unique predictor and language was not, relative to the model without fractions and proportional reasoning). All indirect effects operated via the same variables as the former model; in addition, three small specific indirect effects were present in the latter model but not the former one.

Discussion

The aim of this study was to evaluate domain-specific and domain-general contributions to key mathematical skills in middle school and to situate these within a hierarchy of distal to proximal effects. Both general and specific processes were relevant, although frequently the more proximal arithmetic skills mediate the impact of the general processes and more distal specific processes. Direct effects for the most distal (cognitive) processes were weak and primarily evident only for the general mathematics measure. Overall, a clear hierarchy was evidenced, with sixth-grade math performance relying most heavily on whole number arithmetic skill (procedural computations). The influence of math facts on middle school mathematics was mediated by procedural computations, and in turn the effect of numerosity (particularly symbolic comparison) was mediated by fact retrieval. However, for more complex outcomes, the number of direct effects from cognitive and numerosity domains increased.

Fraction performance

The results for fraction performance were only partially consistent with expectations given that the only significant direct effect on fractions was for procedural computations. Visuospatial and

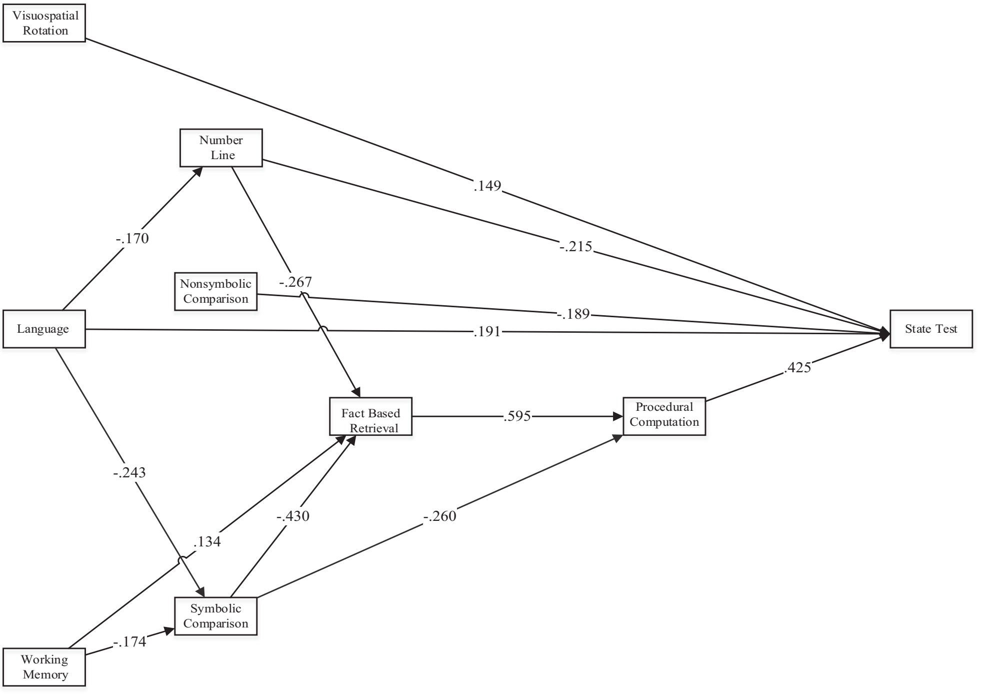

Fig. 3. Path model predicting general mathematics. Values shown are standardized beta weights. Only significant paths are shown for clarity. Standard errors and residuals are not shown for space reasons. Select indirect path statistics (all significant indirect paths and all cumulative indirect effects) appear in Table 3. All specific significant indirect effects in Table 3 are computable from the values in the figure.

nonsymbolic comparison showed neither direct nor indirect effects, but all other (cognitive and numerosity) predictors did show indirect effects. All indirect paths operated through procedural computations; in turn, all paths to procedural computations operated through math facts or symbolic comparison. These findings demonstrate the importance of symbolic representations and computational manipulations for fraction competency in middle school. Results are consistent with recent studies of fraction performance, showing the importance of whole number skills for fraction skill (Hecht & Vagi, 2010; Jordan et al., 2013).

Regarding domain-general predictors, Fennema and Tartre (1985) demonstrated the relation of spatial visualization to fractions, although that study did not control for the range of constructs considered here. On the other hand, Siegler and Pyke (2013) did not show contributions of working memory to eighth-grade fraction competence, although only direct effects were considered and sixth-grade fraction performance was included in their models. That study showed domain-general contributions, although the primary cognitive measure was IQ and behavioral inattention was also included as a predictor. Although one could argue that inclusion of IQ in the current study would yield differing results, the current study did include a vocabulary measure from an IQ battery and it is the most heavily "g-loaded" measure (e.g., Sattler, 2008)[1]; furthermore, Jordan and colleagues (2013) did not show a relation of nonverbal reasoning to procedural fractions. Working memory in Jordan and colleagues' study was not related to conceptual fractions but was related to

[1] We did not have a measure of nonverbal reasoning to form an IQ composite directly. However, we do note that there was no significant difference in the two language measures used here with the outcomes; where this difference was directionally the largest (favoring vocabulary with regard to proportional reasoning), post hoc models employing only the vocabulary measures yielded a highly similar pattern of direct and indirect effects.

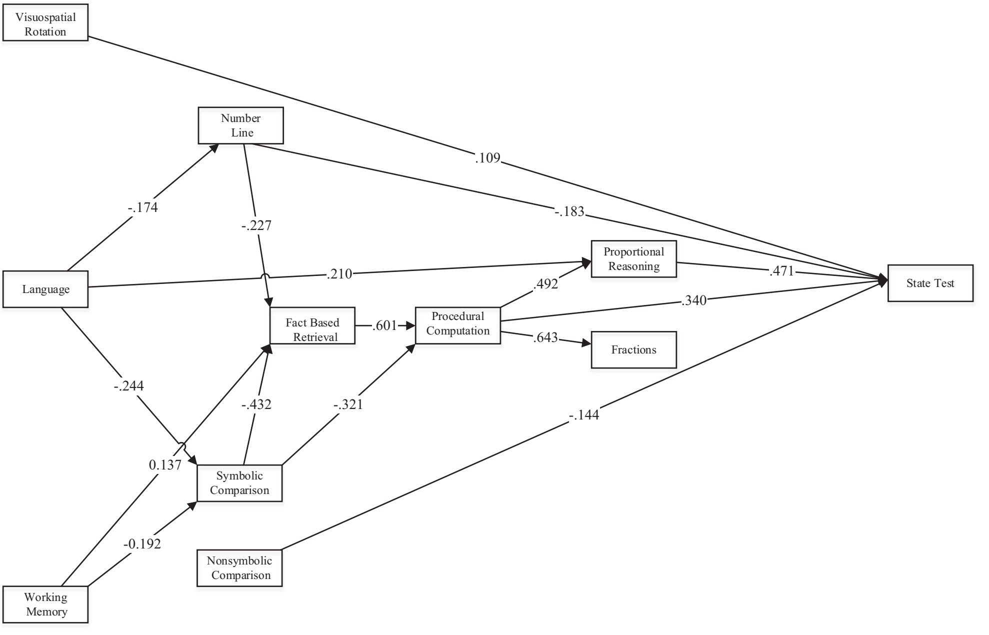

Fig. 4. Path model predicting general mathematics also including fractions and proportional reasoning. Values shown are standardized beta weights. Only significant paths are shown for clarity. Standard errors and residuals are not shown for space reasons. Select indirect path statistics (all significant indirect paths and all cumulative indirect effects) appear in Table 3. All specific significant indirect effects in Table 3 are computable from the values in the figure.

procedural fractions; in the current study, the effects of working memory were only indirect. However, the current results are consistent with those of Vukovic and colleagues (2014), where domain-general predictors were only indirectly related to fractions (concepts).

Regarding numerosity, the only study to evaluate the contribution of nonsymbolic number skills to fractions was Jordan and colleagues (2013), who did not find a relation to either procedural or conceptual fractions, and in that respect results of the current study are consistent. We did not show direct effects of basic facts, whereas Jordan and colleagues did, although that study did not also include procedural computations in its predictive models and both are indicators of whole number facility. Jordan and colleagues showed direct effects for number line estimation, whereas here its effect was only indirect. In Vukovic and colleagues (2014), only a measure of fluency in processing numerical magnitude with both symbolic and nonsymbolic stimuli showed a direct effect across grades; however, it is not clear whether this would be the case in the context of whole number arithmetic.

In sum, the results here are consistent with prior work in showing that fraction performance is related to a wide variety of predictors in both cognitive and numerosity domains. Our results extend these by more strongly implicating whole number facility, particularly when it is more complex than basic facts. More work is needed to more precisely elucidate these relationships when there is an identifiable skill hierarchy present. An important caveat is that our fraction measure was unable to cleanly separate procedural components from conceptual ones, although items of both types were included, whereas other studies included separate measures of procedural and conceptual skills (e.g., Jordan et al., 2013). It could be argued that the role of procedural computations is likely to be larger in the case of computational and procedural fractions relative to those that emphasize conceptual and/or word problems; in the case of the latter, one could argue that more direct contributions from language, working memory, and magnitude understanding in particular would emerge. Future studies would be needed to evaluate these hypotheses directly.

Proportional reasoning

There was strong correspondence in terms of the predictors of fraction performance and proportional reasoning. This is perhaps not surprising given that proportions can be expressed as fractions. In the one study that specifically examined predictors of proportional reasoning, Kwon and colleagues (2000) found working memory to be a significant predictor along with other executive and visuospatial skills. Working memory did not show direct effects in the current study, but it was a contributor to symbolic comparison, which in turn had indirect effects, primarily though procedural computations. The visuospatial measure used in the current study did not contribute to proportional reasoning, whereas the visuoperceptual measure of Kwon and colleagues did. However, results are difficult to compare because it should be noted that, relative to the current study, those students were older, the sample size was smaller, and the students were selected for difficulty in proportional reasoning and received one of two types of proportional reasoning tutoring. Furthermore, the visuoperceptual measure used by Kwon and colleagues was a measure of embedded figures, whereas the current study used a rotational measure.

Some differences relative to fractions prediction were also evident. The current study found direct effects of language on proportional reasoning, consistent with an older study (Lawson, Lawson, & Lawson, 1984) that identified linguistic skills as crucial for solving problems involving proportional reasoning. The linguistic demands of the proportional reasoning measure used here were more prominent than those of the fraction measure. Because students read and answered the proportional reasoning questions on their own, this suggests that both the content and style in which problems are presented, not just the mathematical manipulations, are relevant to determining the corpus of abilities and skills needed for successful performance. This may be more evident at later ages relative to studies of younger students, where problems are often read aloud to students (e.g., Cirino, Fuchs, Elias, Powell, & Schumacher, 2015). The current study also found an indirect role for number line estimation (via fact retrieval), and Rouder and Geary (2014) recently identified a link between number line performance and proportional reasoning. Multidigit skills, such as number line estimation, were also a strong predictor of a variety of mathematical skills, including math facts, written computation, and word problems in Grade 3 (Cowan & Powell, 2014), although that study did not also consider the hierarchy of math skills.

For both fraction performance and proportional reasoning, somewhat surprisingly, domain-general measures of working memory and visuospatial rotation showed only weak effects (directly or indirectly). These findings suggest that for students who struggle with mathematics, it is likely more fruitful to focus on directly teaching lower levels in the math hierarchy than otherwise attempting to increase more removed (cognitive) abilities, particularly at later ages. However, to the extent that foundational skills are weak, and given that indirect effects were clearly present, developing strategies that mitigate cognitive demands on more intermediary skills may have “downstream” implications for later developing skill.

An alternative consideration involves the use of a visuospatial working memory measure in the current study, which could promote overlapping contributions relative to the visuospatial rotation measure as well as diminish the apparent role of working memory. The two measures were significantly (albeit weakly) related in the current study, and each showed similarly modest relationships with the three outcomes and in the case of working memory with whole number arithmetic (see Table 2). However, in post hoc analyses, dropping either of these measures from final models did not change conclusions regarding the other measures’ direct or indirect significance. Given the influence of language for this proportional reasoning measure, as well as the consistent (albeit indirect) role of symbolic comparison of Arabic numerals as well as whole number arithmetic, it is possible that more verbally mediated measures of working memory (e.g., Digit Span, Listening Span, Counting Span) could have yielded stronger ties to the outcomes. We view this as relatively unlikely, at least when language is directly measured and when proximal whole number skills are also included. Although the current results do provide strong initial results for how these sets of predictors impact both fractions and proportional reasoning, more nuanced studies that more carefully parameterize these measures along the lines of content (procedural vs. conceptual) and presentation (e.g., computational, graphical, word problems) are warranted.

State test

The general mathematics measure was substantially broader in content than either of the other intermediary math outcomes (e.g., items involving numerical patterning, spatial reasoning, and measurement were specifically included). Although the proportion of variance accounted for by all domain-general and domain-specific predictors, along with arithmetic, was similar for all three outcomes (range = 38–47% variance), the unique contributors to general mathematics included procedural computations (similar to both fractions and proportional reasoning) as well as language (similar to proportional reasoning). However, there was some divergence, with visuospatial rotation, number line estimation, and nonsymbolic comparison each also having direct effects for this outcome. To a greater degree than either the fraction or proportional reasoning measure, the state test presented most items as word problems, which may more readily engage language skills. Similarly, the role of visuospatial rotation may have been more direct here given that one grouping of the state test items included geometry and spatial reasoning. This pattern was solidified even when fractions and proportional reasoning were included as additional indirect predictors, with the exception that now most indirect effects occurred via proportional reasoning, including the previously direct effect of language. An alternative explanation for the additional contributions of domain-general predictors to the state test may involve the fact that the state test addresses topics in the sixth-grade curriculum. It may be that the contribution of domain-general skills is most relevant during learning. Longitudinal studies would be helpful in this regard.

As with the other mathematical outcomes, somewhat more puzzling was the *lack* of a direct effect for working memory. In the model without fractions and proportional reasoning, there was only a cumulative indirect effect, but with those intermediary outcomes working memory did show several (albeit small) indirect effects through procedural computations and proportional reasoning. Three possibilities may be operating. One is that for very broad tasks (e.g., the state test), the effect of working memory must compete with a wider array of demands (e.g., the language-based presentation, whether graphs or figures are required). Second, the resource demands on working memory operate at lower levels such that they are depleted in the context of nonfluent whole number skill. If this were true, then one might expect that working memory might be a stronger direct predictor of broad mathematical competency when more basic elements of the mathematical hierarchy (and perhaps other cognitive resources) are at least adequate, although this question was beyond the scope of the current study (and perhaps its sample size). Third, working memory was measured with a single task, and more comprehensive measurement may have produced more robust results. We view all of these possibilities as relevant for thinking about the divergence of these results (especially for working memory) relative to prior studies of younger students. The current results do, however, imply that in the context of hierarchical skills, the role of working memory and other predictors needs to consider these more proximal measures.

Among numerosity predictors, symbolic comparison showed only indirect effects, whereas both number line estimation and nonsymbolic comparison showed direct effects. Again, the task items, including measurement, may account in part for the wider array of not only cognitive but also numerosity predictors. Siegler and Lortie-Forgues' (2014) integrative theory of numerical development provides another potential explanation in that a more clear representation of quantity and their manipulation is reintroduced with more complex tasks. However, it is not immediately clear why nonsymbolic comparison in particular would exhibit consistent effects, particularly given the array of other predictors considered in the model, and only for the state test and not fraction performance and proportional reasoning. It is notable that several prominent studies of this effect involved older students (e.g., Halberda et al., 2008; Mazzocco et al., 2011), including a training study in adults (DeWind & Brannon, 2012). One recent meta-analysis did not find age differences in the size of the effect (Chen & Li, 2014), although another one did (Fazio et al., 2014). It is interesting to note that the relationship between nonsymbolic comparison with the fraction, proportional reasoning, and state test outcomes varied slightly (see Table 2; *r*s = .201, .176, and .247, respectively). The directionally highest correlation was for the state test measure, which was slightly higher than the overall size of these meta-analytic effects relevant to this age range (*r*s = .20 and .17, respectively) and may be due in part to the wider range of content and variability for this measure.

That proportional reasoning was more impactful of general mathematics competence than fractions may be in part due to the necessity of identifying relationships among quantities in the service of problem solving, or it may be that the proportional reasoning measure, more so than fractions, was framed in terms of word problems. Either way, considering school-related mathematics, the kinds of skills important for performance included the full range of predictors across cognitive/neuropsychological measures and domain-specific numerosity measures as well as arithmetic and other intermediary math outcomes.

Limitations and conclusions

Before offering conclusions, we draw readers' attention to study limitations. First, although results of the path analyses appear as a step-by-step depiction of how complex math skills develop, the current study is limited to concurrent relations at one grade level. Longitudinal studies are required to establish the true extent to which earlier developing skills are impactful of later developing ones. Even cross-sectional studies across a grade range would add at least indirect evidence of how skills differ at different developmental or learning epochs. Furthermore, the model fits evidenced here would be equal to alternative models with similar paths, albeit in different directions. However, it was the case that the ordering of the variables was based on previous studies indicating the hierarchical and developmental progression of mathematics. The path analytic framework is a useful one for considering direct and indirect effects. Although a more fully latent variable approach may have offered a more complete picture, such measurement would have required a likely prohibitive investment of time and of school and participant resources. This was not possible in the current situation. Clearly, building on these findings in longitudinal fashion would be one way to enhance the impact of these results. At the same time, the current findings do offer testable predictions for how such development might occur and what types of skills ought to emerge in what order.

Furthermore, as noted, outcome measures could have better distinguished between procedural and conceptual outcomes as well as other parameters related to their presentation. Doing so may have revealed different sets of predictors of, for example, fraction concepts versus fraction arithmetic, particularly over time, as others have found (Hecht & Vagi, 2010; Vukovic et al., 2014). Again, this study is one step toward considering a wide array of useful predictors, which when taken together offer information on these three outcomes that may be expanded in future studies.

Third, although a number of predictors were included, not all potential variables were considered. For example, working memory is one primary factor found to be important for math but could have been assessed more comprehensively (e.g., with verbal content), and other skills within the executive function family (e.g., inhibition, shifting, planning, fluency) may have been considered as well. Additional skills such as attention (assessed via behavioral ratings and/or cognitive measures) may also play a predictive role. Within the number domain, assessing other domains of measurement (e.g., fraction number line) or of whole number (e.g., subtraction, division) may have offered a more complete picture.

With these study limitations in mind, we offer the following conclusions. The current study, first and foremost, highlights the progressive development of mathematical skills. It also demonstrates that along a procedural or computational continuum, it is critically important to understand students' whole number arithmetic skills as well as the foundational contributors to these skills, which can be domain general (cognitive) or domain specific (numerosity). These findings are consistent with studies of younger students demonstrating that a wider range of skills is necessary for more complex mathematics (Cowan & Powell, 2014; Fuchs et al., 2010a, 2010b) but extends this pattern to students in middle school. Thus, our hypothesis about both cognitive and numerosity predictors having an impact on sixth-grade mathematics was generally supported. Their role was either direct or indirect (via whole number arithmetic), with direct effects being more common for mathematical skills that require multiple components (e.g., probability, statistics, geometry), differential presentations (e.g., charts, graphs, tables), and increasing language (e.g., word problems).

Acknowledgments

This research was supported by a grant from the Institute of Education Sciences (R305A1100067, Arithmetical and Cognitive Antecedents and Concomitants of Algebraic Skill) awarded to the University of Houston. The attitudes and opinions expressed in this presentation are those of the authors and do not necessarily reflect those of the funding agency. The authors thank the colleagues, parents, teachers, and school and district officials who made this research possible.

References

Alloway, T., Gathercole, S., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development, 80*, 606–621.
Ashcraft, M. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44*, 75–106.
Ashcraft, M., & Christy, K. (1995). The frequency of arithmetic facts in elementary texts: Addition and multiplication in Grades 1–6. *Journal for Research in Mathematics Education, 26*, 396–421.
Aunio, P., & Niemivirta, M. (2010). Predicting children's mathematical performance in Grade One by early numeracy. *Learning and Individual Differences, 20*, 427–435.
Aunola, K., Leskinen, E., Lerkkanen, M., & Nurmi, J. (2004). Developmental dynamics of math performance from preschool to Grade 2. *Journal of Educational Psychology, 96*, 699–713.
Bailey, D., Hoard, M., Nugent, L., & Geary, D. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology, 113*, 447–455.
Booth, J., Newton, K., & Twiss-Garrity, L. (2014). The impact of fraction magnitude knowledge on algebra performance and learning. *Journal of Experimental Child Psychology, 118*, 110–118.
Booth, J., & Siegler, R. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 41*, 189–201.
Brown, G., & Quinn, R. (2006). Algebra students' difficulty with fractions: An error analysis. *The Australian Mathematics Teacher, 62*(4), 28–40.
Brown, G., & Quinn, R. J. (2007a). Fraction proficiency and success in algebra: What does the research say? *The Australian Mathematics Teacher, 63*(3), 23–30.
Brown, G., & Quinn, R. (2007b). Investigating the relationship between fraction proficiency and success in algebra. *The Australian Mathematics Teacher, 63*(4), 8–15.
Butterworth, B. (1999). *The mathematical brain*. London: Macmillan.
Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry, 46*, 3–18.
Bynner, J. (2012). Policy reflections guided by longitudinal study, youth training, social exclusion, and more recently Neet. *British Journal of Educational Studies, 60*, 39–52.
Calhoon, M., Emerson, R., Flores, M., & Houchins, D. (2007). Computational fluency performance profile of high school students with mathematics disabilities. *Remedial and Special Education, 28*, 292–303.
Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica, 148*, 163–172.
Cirino, P. (2011). The interrelationships of mathematical precursors in kindergarten. *Journal of Experimental Child Psychology, 108*, 713–733.
Cirino, P., Fuchs, L., Elias, J., Powell, S., & Schumacher, R. (2015). Cognitive and mathematical profiles for different forms of learning difficulties. *Journal of Learning Disabilities, 48*, 156–175.
Cirino, P., Morris, M., & Morris, R. (2002). Neuropsychological concomitants of calculation skills in college students referred for learning disabilities. *Developmental Neuropsychology, 21*, 201–218.
Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*, 769–786.
Cowan, R., & Powell, D. (2014). The contributions of domain-general and numerical factors to third-grade arithmetic skills and mathematical learning disability. *Journal of Educational Psychology, 106*, 214–229.
Dehaene, S. (2001). Précis of the number sense. *Mind & Language, 16*, 16–36.
Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition, 1*, 83–120.
Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*, 487–506.
DeWind, N., & Brannon, E. (2012). Malleability of the approximate number system: Effects of feedback and training. *Frontiers in Human Neuroscience, 6*. http://dx.doi.org/10.3389/fnhum.2012.00068.
Ekstrom, R. B., French, J. W., Harman, H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests* (rev. ed.). Princeton, NJ: Educational Testing Service.
Fazio, L., Bailey, D., Thompson, C., & Siegler, R. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology, 123*, 53–72.
Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*, 307–314.
Fennema, E., & Tartre, L. (1985). The use of spatial visualization in mathematics by girls and boys. *Journal for Research in Mathematics Education, 16*, 184–206.
Fuchs, L., Compton, D., Fuchs, D., Paulsen, K., Bryant, J., & Hamlett, C. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology, 97*, 493–513.
Fuchs, L., Fuchs, D., Stuebing, K., Fletcher, J., Hamlett, C., & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology, 100*, 30–47.

Fuchs, L., Geary, D., Compton, D., Fuchs, D., Hamlett, C., & Bryant, J. (2010a). The contributions of numerosity and domain-general abilities to school readiness. *Child Development, 81*, 1520–1533.
Fuchs, L., Geary, D., Compton, D., Fuchs, D., Hamlett, C., Seethaler, P., ... Schatschneider, C. (2010b). Do different types of school mathematics development depend on different constellations of numerical versus general cognitive abilities? *Developmental Psychology, 46*, 1731–1746.
Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C. L., Cirino, P. T., ... Changas, P. (2013). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology, 105*(3), 683–700.
Fuchs, L., Geary, D., Fuchs, D., Compton, D., & Hamlett, C. (2014). Sources of individual differences in emerging competence with numeration understanding versus multidigit calculation skill. *Journal of Educational Psychology, 106*, 482–498.
Gabel, D. L., Sherwood, R. D., & Enochs, L. (1984). Problem-solving skills of high school chemistry students. *Journal of Research in Science Teaching, 21*, 221–233.
Gabriel, F., Coché, F., Szucs, D., Carette, V., Rey, B., & Content, A. (2012). Developing children's understanding of fractions: An intervention study. *Mind, Brain, and Education, 6*(3), 137–146.
Gathercole, S., Brown, L., & Pickering, S. (2003). Working memory assessments at school entry as longitudinal predictors of national curriculum attainment levels. *Educational and Child Psychology, 20*(3), 109–122.
Geary, D. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114*, 345–362.
Geary, D. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37*, 4–15.
Geary, D. (2013). Early foundations of mathematics learning and their relations to learning disabilities. *Current Directions in Psychological Science, 22*, 23–27.
Geary, D. C., Bow-Thomas, C. C., Liu, F., & Siegler, R. S. (1996). Development and arithmetic competencies in Chinese and American children: Influence of age, language, and schooling. *Child Development, 67*, 2022–2044.
Geary, D., Hoard, M., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development, 78*, 1343–1359.
Geary, D., Hoard, M., Nugent, L., & Rouder, J. (2015). Individual differences in algebraic cognition: Relation to the approximate number and semantic memory systems. *Journal of Experimental Child Psychology, 140*, 211–227.
Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in nonverbal number acuity predict maths achievement. *Nature, 455*, 665–668.
Hallett, D., Nunes, T., & Bryant, P. (2010). Individual differences in conceptual and procedural knowledge when learning fractions. *Journal of Educational Psychology, 102*, 395–406.
Hayes, A. (2009). Beyond Baron and Kenny: Statistical mediation analysis in new millennium. *Communication Monographs, 76*, 408–420.
Hecht, S., & Vagi, K. (2010). Sources of group and individual differences in emerging fraction skills. *Journal of Educational Psychology, 102*, 843–859.
Hecht, S., Vagi, K., & Torgesen, J. (2007). Fraction skills and proportional reasoning. In D. B. Berch & M. M. M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 121–132). Baltimore, MD: Paul H. Brookes.
Holloway, I., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology, 103*, 17–29.
Hyde, D., Khanum, S., & Spelke, E. (2014). Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children. *Cognition, 131*, 92–107.
Jordan, N., Hanich, L., & Kaplan, D. (2003). A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with comorbid mathematics and reading difficulties. *Child Development, 74*, 834–850.
Jordan, N., Hansen, N., Fuchs, L., Siegler, R., Gersten, R., & Micklos, D. (2013). Developmental predictors of fraction concepts and procedures. *Journal of Experimental Child Psychology, 116*, 45–58.
Jordan, N., Kaplan, D., Ramineni, C., & Locuniak, M. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45*, 850–867.
Kane, M., Hambrick, D., Tuholski, S., Wilhelm, O., Payne, T., & Engle, R. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*, 189–217.
Kilpatrick, J., Swafford, J., & Findell, B. (2001, January 1). *Adding it up: Helping children learn mathematics.* Retrieved December 10, 2014, from <http://www.nap.edu/catalog/9822.html>.
Kingston, J., & Lyddy, F. (2013). Self-efficacy and short-term memory capacity as predictors of proportional reasoning. *Learning and Individual Differences, 26*, 185–190.
Klein, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford.
Koponen, T., Aunola, K., Ahonen, T., & Nurmi, J. (2007). Cognitive predictors of single-digit and procedural calculation skills and their covariation with reading skill. *Journal of Experimental Child Psychology, 97*, 220–241.
Koponen, T., Mononen, R., Rasanen, P., & Ahonen, T. (2006). Basic numeracy in children with specific language impairment: Heterogeneity and connections to language. *Journal of Speech, Language, and Hearing Research, 49*, 58–73.
Koponen, T., Salmi, P., Eklund, K., & Aro, T. (2013). Counting and RAN: Predictors of arithmetic calculation and reading fluency. *Journal of Educational Psychology, 105*, 162–175.
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.
Kwon, Y., Lawson, A., Chung, W., & Kim, Y. (2000). Effect on development of proportional reasoning skill of physical experience and cognitive abilities associated with prefrontal lobe activity. *Journal of Research in Science Teaching, 37*, 1171–1181.
Lawson, A. E., Lawson, D. I., & Lawson, C. A. (1984). Proportional reasoning and the linguistic abilities required for hypothetico-deductive reasoning. *Journal of Research in Science Teaching, 21*, 119–131.

Lee, K., Ng, E., & Ng, S. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology, 101*, 373–387.

LeFevre, J., Fast, L., Skwarchuk, S., Smith-Chant, B., Bisanz, J., Kamawar, D., et al (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development, 81*, 1753–1767.

Libertus, M., Odic, D., & Halberda, J. (2012). Intuitive sense of number correlates with math scores on college-entrance examination. *Acta Psychologica, 141*, 373–379.

Lyons, I., Ansari, D., & Beilock, S. (2012). Symbolic estrangement: Evidence against a strong association between numerical symbols and the quantities they represent. *Journal of Experimental Psychology: General, 141*, 635–641.

Lyons, I., & Beilock, S. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition, 121*, 256–261.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Mahwah, NJ: Lawrence Erlbaum.

MacKinnon, D. P., & Fairchild, A. J. (2009). Current directions in mediation analysis. *Current Directions in Psychological Science, 18*, 16–20.

Martin, R., Cirino, P., Charp, C., & Barnes, M. (2014). Number and counting skills in kindergarten as predictors of Grade 1 mathematical skills. *Learning and Individual Differences, 34*, 12–23.

Matteson, S. (2006). Mathematical literacy and standardized mathematical assessments. *Reading Psychology, 27*, 205–233.

Mazzocco, M., Feigenson, L., & Halberda, J. (2011). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development, 82*, 1224–1237.

McGrew, K. S., & Woodcock, R. W. (2001). *Woodcock-Johnson III: Technical manual*. Itasca, IL: Riverside.

Misailidou, C., & Williams, J. (2003). Diagnostic assessment of children's proportional reasoning. *Journal of Mathematical Behavior, 22*, 335–368.

Mundy, E., & Gilmore, C. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology, 103*, 490–502.

Muthen, L. K., & Muthen, B. O. (2007). *Mplus user's guide: Statistical analysis with latent variables*. Los Angeles: Muthen & Muthen.

National Mathematics Advisory Panel (2008). *Reports of the task groups and subcommittees*. Washington, DC: Author. http://www2.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf.

Noël, M., & Rousselle, L. (2011). Developmental changes in the profiles of dyscalculia: An explanation based on a double exact-and-approximate number representation model. *Frontiers in Human Neuroscience, 5*. http://dx.doi.org/10.3389/fnhum.2011.00165.

Parsons, S., & Bynner, J. (2008, February 1). *Illuminating disadvantage: Profiling the experiences of adults with entry level literacy or numeracy over the life course.* Retrieved December 1, 2014, from <http://www.nrdc.org.uk/>.

Peters, M. (1995). *Revised Vandenberg & Kuse mental rotations tests: Forms MRT-A to MRT-D*. Guelph, Ontario, Canada: University of Guelph, Department of Psychology.

Preacher, K., & Hayes, A. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*, 717–731.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879–891.

Raghubar, K., Barnes, M., & Hecht, S. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*, 110–122.

Rakes, C. R., Valentine, J. C., McGatha, M. B., & Ronau, R. N. (2010). Methods of instructional improvement in algebra: A systematic review and meta-analysis. *Review of Educational Research, 80*(3), 372–400.

Rose, H., & Betts, J. (2004). The effect of high school courses on earnings. *Review of Economics and Statistics, 86*, 497–513.

Rouder, J., & Geary, D. (2014). Children's cognitive representation of the mathematical number line. *Developmental Science, 17*, 525–536.

Rousselle, L., & Noël, M. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs. non-symbolic number magnitude processing. *Cognition, 102*, 361–395.

Sattler, J. (2008). *Assessment of children: Cognitive foundations* (5th ed.). San Diego: J. M. Sattler.

Shrout, P., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*, 422–445.

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

Siegler, R., Duncan, G., Davis-Kean, P., Duckworth, K., Claessens, A., Engel, M., ... Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*, 691–697.

Siegler, R., Carpenter, T., Fennell, F., Geary, D., Lewis, J., Okamoto, Y., . . .Wray, J. (2010). *Developing effective fractions instruction for kindergarten through 8th grade: A practice guide* (NCEE, No. 2010-4039). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved on April 17, 2015, from <http://www.whatworks.ed.gov/publications/practiceguides>.

Siegler, R., Fazio, L., Bailey, D., & Zhou, X. (2013). Fractions: The new frontier for theories of numerical development. *Trends in Cognitive Sciences, 17*, 13–19.

Siegler, R., & Lortie-Forgues, H. (2014). An integrative theory of numerical development. *Child Development Perspectives, 8*(3), 144–150.

Siegler, R., & Opfer, J. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*, 237–250.

Siegler, R., & Pyke, A. (2013). Developmental and individual differences in understanding of fractions. *Developmental Psychology, 49*, 1994–2004.

Siegler, R., & Ramani, G. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology, 101*, 545–560.

Siegler, R., Thompson, C., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62*, 273–296.

Szucs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2014). Cognitive components of a mathematical processing network in 9-year-old children. *Developmental Science, 17*, 506–524.

Tolar, T., Lederberg, A., & Fletcher, J. (2009). A structural model of algebra achievement: Computational fluency and spatial visualization as mediators of the effect of working memory on algebra achievement. *Educational Psychology, 29*, 239–266.
Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory, 17*, 635–654.
Vandenberg, S., & Kuse, A. (1978). Mental rotation, a group test of 3-D spatial visualization. *Perceptual and Motor Skills, 47*, 599–604.
vanMarle, K., Chu, F. W., Li, Y., & Geary, D. C. (2014). Acuity of the approximate number system and preschoolers' quantitative development. *Developmental Science, 17*, 492–505.
Vogel, C. (2008, May 1). Algebra: Changing the equation: Support programs decrease math anxiety and help students succeed. *District Administration,* pp. 34–40. Available at <http://www.districtadministration.com/article/algebra-changing-equation>.
von Aster, M., & Shalev, R. (2007). Number development and developmental dyscalculia. *Developmental Medicine and Child Neurology, 49*, 868–873.
Vukovic, R., Fuchs, L., Geary, D., Jordan, N., Gersten, R., & Siegler, R. (2014). Sources of individual differences in children's understanding of fractions. *Child Development, 85*, 1461–1476.
Vukovic, R., & Lesaux, N. (2013). The language of mathematics: Investigating the ways language counts for children's mathematical development. *Journal of Experimental Child Psychology, 115*, 227–244.
Wechsler, D. (1999). *Wechsler Abbreviated Scales of Intelligence: WASI*. San Antonio, TX: Psychological Corporation.
Woodcock, R., McGrew, K., & Mather, N. (2001). *Woodcock-Johnson III tests of achievement*. Itasca, IL: Riverside.