Published in final edited form as:
Child Dev. 2010 ; 81(5): 1520–1533. doi:10.1111/j.1467-8624.2010.01489.x.

The Contributions of Numerosity and Domain-General Abilities to School Readiness

Lynn S. Fuchs,
Vanderbilt University

David C. Geary,
University of Missouri

Donald L. Compton,
Vanderbilt University

Douglas Fuchs,
Vanderbilt University

Carol L. Hamlett, and
Vanderbilt University

Joan D. Bryant
Vanderbilt University

Abstract

Contributions of domain-general and domain-specific numerical competencies were assessed on 1st graders' number combination skill (NC) and word problem skill (WP). Students (*n*=205) between 5–7 years of age were assessed on 2 aspects of numerosity, 8 domain-general abilities, NC, and WP. Both aspects of numerosity predicted NC when controlling for domain-general abilities, but domain-general abilities did not account for significant additional variance. By contrast, when controlling for domain-general abilities in predicting WP, only precise representation of small quantities was uniquely predictive, and domain-general measures accounted for significant additional variance; central executive component of working memory and concept formation were uniquely predictive. Results suggest that development of NC and WP depends on different constellations of numerical versus more general cognitive abilities.

In an analysis of six large-scale longitudinal studies, Duncan et al. (2008) demonstrated that mathematical competence at school entry predicts mathematics achievement throughout the elementary-school years, above and beyond general cognitive ability, classroom attention, social skills, or socioeconomic background. In fact, performance on early mathematical achievement tests was by far the single best predictor of later mathematics achievement. They suggested "it may be beneficial to add domain-specific early skills to the definition of school readiness" (Duncan et al., p. 1429), but their analysis did not allow for the assessment of which specific mathematical competencies may be the best target for such programs. Candidates center on children's early number sense, including the ability to quickly apprehend the quantities of small sets of items, use counting to determine quantity, estimate the value of large quantities, and intuitively understand the effects of addition and subtraction on quantity (National Mathematics Advisory Panel, 2008).

Inquiries should be sent to Lynn S. Fuchs, 228 Peabody, Vanderbilt University, Nashville, TN 37203.

At the same time, Duncan et al. (2008) showed that classroom attention, a domain-general factor, also predicts later mathematics achievement, above and beyond early mathematical competence. Other studies have revealed that general cognitive ability is also a strong predictor of achievement across academic domains (e.g., Walberg, 1984). General cognitive ability includes working memory capacity, speed of information processing, and logical reasoning (Embretson, 1995; Engle, Tuholski, Laughlin, & Conway, 1999; Kail, 1991), although the relative importance of these domain-general abilities is debated (e.g., Ackerman, Beier, & Boyle, 2005).

In the present study, we focused on the relation between two measures of children's early number sense as well as domain-general cognitive and attentional measures with performance on two critical mathematical abilities: fluency in solving arithmetic problems (e.g., 5+7=12; 9−5=4), which are sometimes referred to as number combinations (NCs; Baroody, 1985), and simple word problems (WPs) that require solution of NCs. By assessing competencies in NC and WP, rather than using broad achievement measures, we explored the relative importance of domain-specific versus domain-general abilities to each of these core aspects of children's early mathematical learning. If Duncan et al. (2008) are correct, then measures of children's number sense should predict NC and WP competence, with domain-general measures accounting for relatively little additional variance. Such findings would help to more firmly establish links between children's informal (before schooling) number sense and their emerging competencies in school-taught mathematics and provide direction for the foci of school readiness programs.

The Domain-General Mechanisms

Of the domain-general mechanisms, working memory, classroom attention, and speed of processing have all been show to predict NC performance (e.g., Bull & Johnston, 1997; Fuchs et al., 2005; Hitch & McAuley, 1991), but none of the studies assessed these potential mechanisms simultaneously with each other and with domain-specific measures of number sense. Moreover, working memory itself constitutes multiple sub-competencies that may influence NC development in different ways.

At the least, working memory comprises an attention-driven central executive, which controls the representation and manipulation of information in two sub-systems that provide short-term storage (Baddeley, 1986). The phonological loop stores auditory or phonological information; the visuospatial sketchpad stores spatial and visual information such as shapes and colors or the location or speed of objects in space. Correlational studies suggest the central executive is critical for NC development, (e.g., Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Swanson & Sachse-Lee, 2001), but the subcomponents of the central executive contributing to this performance, including attentional and inhibitory control, are not well understood. Attentional control may be necessary to maintain the simultaneous activation of problem stems and answers in working memory while counting solutions; inhibitory control may be needed to prevent irrelevant associations from intruding into working memory during problem solving (Barrouillet, Fayol, & Lathuliere, 1997; Engle et al., 1999).

The phonological loop may contribute to the establishment of NCs in long-term memory, because children must encode and maintain accurate phonological representations of problem stems and answers in working memory as they count to solve problems (e.g., Logie & Baddeley, 1987). Evidence on the link between phonological processing and NC skill is, however, inconsistent (e.g., Fuchs et al., 2005; Hecht, Torgesen, Wagner, & Rashotte, 2001; Swanson & Beebe-Frankenberger, 2004). Attentive behavior in the classroom can also influence NC skill (Fuchs et al., 2005; Fuchs et al., 2006), but the relation between this type of attention and the attentional and inhibitory control components of the central executive are not well understood.

The search for the mechanisms underlying NC skill is further complicated by the relation between working memory and speed of processing (the efficiency with which simple tasks are executed; Case, 1985). Bull and Johnston (1997) found that behavioral measurement of processing speed was the best predictor of NC skill among 7-year olds, subsuming all the variance accounted for by long- and short-term memory, even with reading skill controlled. Hecht et al. (2001) provided corroborating data while controlling for oral vocabulary, and profile analysis (Fuchs, Fuchs, Stuebing et al., 2008) showed that a behavioral measure of processing speed distinguishes third graders with specific NC deficits from those with WP deficits. However, it is not known whether individual differences in working memory are driven by more fundamental differences in speed of processing (Kail, 1991) or whether the attentional focus associated with the central executive speeds information processing (Engle et al., 1999).

Working memory is also related to WP performance but the contributions of the different components of working memory are not fully understood (e.g., Hitch & McAuley, 1991; Swanson & Sachse-Lee, 2001). Moreover, as noted, performance on working memory measures is correlated with logical reasoning, but processes are distinct to each (Embretson, 1995). General reasoning ability predicts third graders' WP performance even after controlling for working memory (Fuchs et al., 2005; Fuchs et al., 2006; Swanson & Sachse-Lee, 2001). Finally, we must consider oral language comprehension as a contributor to WP skill, above and beyond the contributions of the phonological loop. Jordan, Levine, and Huttenlocher (1995) demonstrated that kindergarten and first-grade children with oral language impairment perform lower than unimpaired peers on WPs. Fuchs et al. (2006) showed the relation between language and WP skill at third grade while controlling for a variety of other variables, and profile analysis (Fuchs, Fuchs, Stuebing et al., 2008) revealed that oral language comprehension distinguishes third graders with WP deficits versus those with specific calculation deficits.

The Domain-Specific Mechanisms

In addition to general cognitive abilities, which apply across academic domains, children may have specialized numerical representations and competence in numerical processing (e.g., Baroody, Bajwa, & Eiland, 2009; Feigenson, Dehaene, & Spelke, 2004). Cognitive (Gilmore, McCarthy, & Spelke, 2007), neuropsychological (Domahs & Delozier, 2005), and brain-imaging (Cantlon, Platt, & Brannon, 2009) studies indicate that number sense mechanisms support children's early learning of school-taught mathematics. As per Feigenson et al. (2004), two core systems, one for exact representation of small quantities and one for approximating larger quantities, may be involved in representing numerical value, although controversy exists about whether these systems are supported by distinct cognitive and brain mechanisms (e.g., Cantlon et al.; Gallistel & Gelman, 1992; Mix, Huttenlocher, & Levine, 2002; vanMarle & Wynn, 2009). Whether the underlying mechanisms are cognitively or neurally distinct, they are functionally so; people of any age can rapidly distinguish the quantities of sets of less than four objects but must count or approximate the values of larger sets and Arabic numerals.

The hypothesized first system supports analog representations of approximate numerical magnitudes (Gallistel & Gelman, 1992; Xu, 2003) and can be assessed based on children's placement of numerals on a number line (e.g., Siegler & Booth, 2004). Making placements on a number line, which are based on the use of this approximate magnitude system, results in a pattern that conforms to the natural logarithm (*Ln*) of the numbers (Feigenson et al., 2004; Gallistel & Gelman, 1992; Siegler & Opfer, 2003). In other words, use of this representational system results in placements that are compressed for larger magnitudes such that the perceived distance between 52 and 53 is smaller than the perceived distance between 2 and 3. With schooling children learn to make placements that conform to the mathematical number line; that is, their placements suggest they understand that the difference between two consecutive numbers is identical regardless of position on the line.

The second hypothesized system enables precise representations of small quantities, the fast apprehension of which is termed *subitizing* (Mandler & Shebo, 1982). With a verbal subitizing task, which involved name identity, Koontz and Berch (1996) asked third and fourth graders with and without a mathematical learning difficulty to determine if combinations of Arabic numerals (e.g., 3-2), number sets (e.g., ■■-■■), or numerals and sets were the same (e.g., 2-■■) or different (e.g. 3-■■). Reaction times for children without mathematics difficulty indicated fast access to representations for quantities of two and three, regardless of whether the code was an Arabic numerical or number set. Children with mathematics difficulty showed fast access to numerosity representations for the quantity of two, but appeared to rely on counting to determine quantities of three. Geary et al. (2007; Geary, Bailey, & Hoard, 2009) developed the *Number Sets Test* to assess related, basic numerical fluency. Children are asked to combine pairs or triplets of Arabic numerals (e.g., 3 4) or sets of objects (e.g., ●● ■■■) and quickly determine if they match a target number (e.g., 5). Verbal subitizing can contribute to fluency if the quantity of small sets of objects is apprehended and then added together (e.g., ●● ■■■ = 5; Geary & Lin, 1998). Children with mathematics difficulty perform poorly on this task compared to peers, independent of general cognitive ability, but performance may be mediated by visuospatial working memory (Geary et al., 2008).

Purpose of the Present Study

In the present study, we assessed the relative contributions of domain-general working memory, attentional, language, and reasoning abilities and domain-specific numerical competencies on individual differences in first graders’ NC and WP performance. We had three goals. The first was to determine if measures of the approximate and exact number representation systems directly affect NC skill when controlling for domain-general abilities. Evidence for one or the other numerosity measure would support Duncan et al.’s (2008) suggestion that domain-specific competence may be important for children’s early mathematical learning. Such a finding would also provide much needed information on where to target early interventions for children at risk for poor outcomes in mathematics (e.g., Siegler & Ramani, 2008). Our second goal was to determine how much of the variance in domain-general abilities and numerosity variables represents unique versus shared variance in explaining NC skill. Finding that domain-general abilities uniquely account for much of the variance would suggest that children’s early mathematics learning is supported by the same mechanisms that support learning in other domains; finding that the numerosity variables uniquely account for much of the variance would substantiate a domain-specific perspective. By contrast, finding that a substantial amount of the variance is shared would suggest the viability of a combined perspective. Our third goal was to address the same issues as in our first two goals but applied to WPs that require solutions to NCs. These analyses assess whether domain-general or domain-specific relations found for NCs generalize to mathematics content beyond NC skill.

Method

Participants

As part of a prospective 4-year study of the development of NC skill beginning in first grade, data for the present study were collected with the first-year sample at the first assessment wave, when we sampled children from 52 first-grade classrooms in nine Title 1 schools and two non-Title 1 schools (2–9 teachers per school) in a southeastern metropolitan school district. From 898 students in these classrooms, we secured consent on 580 (64.6%), who were then screened for inclusion using a 2-step process. Step 1 involved whole-class screening on the First-Grade Test of Computational Fluency and the First-Grade Test of Mathematics Concepts and Applications (Fuchs, Hamlett, & Hamlett, 1990; see Measures). Teachers first used scripts to administer two practice alternate forms of each measure. Prior to actual screening, teachers returned the unscored practice tests to project staff. Screening occurred with one additional alternate form of each measure. In Step 2, all students who were present on the day of whole-class screening (n = 507) were tested individually on measures of early reading: Rapid Letter Naming (Fuchs et al., 2001) and Word Identification Fluency (Fuchs, Fuchs, & Compton, 2004; see Measures). We indexed early reading for screening because evidence indicates that students with concurrent difficulty with mathematics and reading may experience more severe difficulty with NCs (e.g., Hanich, Jordan, Kaplan, & Dick, 2001).

Using mixture analysis that combined the four screening scores into a latent variable, we created high, average, and at-risk strata. We combined the high and low strata into a single “not at risk” group. We then applied the following exclusions. First, we excluded 73 students whose mathematics performance was inconsistent with their overall risk status (i.e., at-risk students with high mathematics scores and not-at-risk students with low mathematics scores). Second, we excluded 121 students whose teachers identified them as not speaking English. Third, we administered the 2-subtest Wechsler Abbreviated Intelligence Scale (WASI; Wechsler, 1999) and excluded seven students with standard scores below 80 on both subtests (i.e., our goal was not to study mental retardation).

From the remaining 306 students, we randomly sampled 240 students, purposely stratifying by the at-risk and not-at-risk strata and ensuring that no more than eight students came from the same classroom. The present sample comprises the 205 students on whom we had complete data on the study variables. These students had mean standard scores in the normal range on the Reading subtest of the Wide Range Achievement Test–3 (WRAT, Wilkinson, 1993), 105.12 (*SD* = 15.35); WRAT-Arithmetic, 96.63 (*SD* = 14.63); KeyMath-Revised (Connolly, 1998)-Numeration, 102.80 (*SD* = 11.29); KeyMath Revised-Addition, 99.90 (*SD* = 10.46); KeyMath Revised-Subtraction, 97.30 (*SD* = 11.97); and WASI IQ, 92.70 (*SD* = 13.31). Of the 205 participants, 109 (53.2%) were boys; 161 (78.5%) received subsidized lunch. Among the students, 112 (54.6%) were African American, 48 (23.4%) were European American, 24 (11.7%) were Hispanic, and 21 (10.2%) were other. Students with a school-identified disability comprised 4.8% of the sample.

Screening Measures to Constitute a Representative Sample

Mathematics screeners—Because these measures were used to screen overall mathematics competence, they included problems with a range of difficulty to be sensitive to the skills of low, average, and high performers.

The *First-Grade Test of Computational Fluency* (Fuchs et al., 1990) is a 1-page testdisplaying 25 items that sample the range of problems typically found in the first-grade computation curriculum: adding two single-digit numbers (9 items), subtracting two single-digit numbers (10 items), adding three single-digit numbers (2 items), adding two 2-digit numbers without regrouping (2 items), and subtracting a 1-digit number from a 2-digit number (2 items). Students have 2 min to complete as many items as possible. The score is the number of digits correct (starting from the right, credit is awarded if the correct numeral is in the correct place so that 38+26=514 would earn 1 digit correct to reflect incorrect regrouping). Staff entered responses into a computerized program on an item-by-item basis, with 15% of tests re-entered by an independent scorer. Data-entry agreement was 99.8% (i.e., number of agreements divided by number of agreement plus disagreement). Coefficient alpha was .97.

The *First-Grade Test of Mathematics Concepts and Applications* (Fuchs et al., 1990) is a 3-page test with 25 items sampling the typical first-grade concepts/applications curriculum (i.e., numeration, concepts, geometry, measurement, applied computation, money, charts/graphs, WPs). For example, items require students to count arrays of items, write the number after, identify smaller/larger quantities, read clocks, and identify operators within number sentences. The tester reads the words in each item aloud. For 20 items, students have 15 sec to respond; for the remaining 5 items, students have 30 sec. The score is the number of correct answers. Staff entered responses into a computerized program on an item-by-item basis, with 15% re-entered by an independent scorer. Data-entry agreement was 98.8%. Coefficient alpha was .93.

Reading—With *Rapid Letter Naming* (D. Fuchs et al., 2001), students are shown a page with 52 letters (26 letters in upper and lowercase) displayed in random order and have 1 min to say letter names. The score is the number of correct letters. If the child finishes before 1 min, the score is prorated. Two-week stability is .94 (D. Fuchs et al.). With *Word Identification Fluency* (Fuchs et al., 2004), students have 1 min to read a list of 50 words randomly sampled from a pool of 100 high-frequency pre-primer, primer, and 1st-grade words. If a student completes reading before 1 min, the score is prorated. We administered two alternate forms and averaged the scores. Alternate-form reliability/stability is .97, and correlations with fall and spring Woodcock Reading Mastery Test-Word Identification (Woodcock, 1998) are .77–.82.

Intelligence—*WASI* (Wechsler, 1999) measures generalized cognitive ability with *Vocabulary* and *Matrix* Reasoning (see description below). Subtest scores are combined to yield an Estimated Full Scale IQ score. Internal consistency reliability exceeds .92.

Domain-General Predictor Measures

Language—We used two tests of language. *Woodcock Diagnostic Reading Battery* (*WDRB) - Listening Comprehension* (Woodcock, 1997) measures the ability to understand sentences or passages. With 38 items, students supply the word missing at the end of each sentence or passage. The test begins with simple verbal analogies and associations and progresses to comprehension involving the ability to discern implications. Testing is discontinued after six consecutive errors. The score is the number of correct responses. Internal consistency reliability is .80 at ages 5–18; the correlation with the WJ-R is .73. *WASI Vocabulary* (Wechsler, 1999) measures expressive vocabulary, verbal knowledge, and foundation of information with 42 items. The first four items present pictures; the student identifies the object in the picture. For the remaining items, the tester says a word that the student defines. Responses are awarded a score of 0, 1, or 2 depending on quality. Testing is discontinued after five consecutive scores of 0. The score is the total number of points. As reported by Zhu (1999), split-half reliability is .86–.87 at ages 6–7; the correlation with the WISC-III Full Scale IQ is .72.

Concept formation—*Woodcock-Johnson Psycho-Educational Battery–Revised* (WJ III)-*Concept Formation* (Woodcock, McGrew, & Mather, 2001), 2001) asks students to identify the rules for concepts when shown illustrations of instances and non-instances of the concept. Examinees earn credit by correctly identifying the rule that governs each concept. Cut-off points determine the ceiling. The score is the number of correct responses. As reported by the test developer, internal consistency reliability is .93.

Nonverbal problem solving—*WASI Matrix Reasoning* (Wechsler, 1999) measures nonverbal reasoning with four types of tasks: pattern completion, classification, analogy, and serial reasoning. The child looks at a matrix from which a section is missing and complete it by saying the number of or pointing to one of five response options. Examinees earn points by identifying the correct missing piece of the matrix. Testing is discontinued after 4 errors on 5 consecutive items or 4 consecutive errors. The score is the number of correct responses. Internal consistency reliability is .94; the correlation with the WISC-III Full Scale IQ is .66.

Attentive behavior—The *SWAN* is an 18-item teacher rating scale (Swanson, 2004), sampling items from the *Diagnostic and Statistical Manual of Mental Disorders-IV* (APA, 1994) criteria for Attention-Deficit/Hyperactivity Disorder for inattention (largely distractibility) (items 1–9) and hyperactivity/impulsivity (items 10–18). Items are rated as 1=Far Below, 2=Below, 3=Slightly Below; 4=Average, 5=Slightly Above, 6=Above, 7=Far Above. We report data for the attentive behavior subscale as the average rating across the nine relevant items. We selected this subscale to operationalize attentive behavior, or the ability to maintain focus of attention. The SWAN correlates well with other dimensional assessments of behavior related to attention (www.adhd.net). Coefficient alpha in the present study was .97.

Working memory—The *Working Memory Test Battery for Children* (WMTB-C; Pickering & Gathercole, 2001) comprises nine subtests that assess the central executive, phonological loop, and visuospatial sketchpad. Each subtest has six items at span levels ranging from 1–6 to 1–9. Passing four items at a level moves the child to the next level. At each span level, the number of items to be remembered increases by one. Failing three items terminates the subtest. The order of subtests is designed to avoid over-taxing any one component area of working memory and is generally arranged from easiest to hardest: Digit Recall, Word List Matching, Word List Recall, Nonword List Recall, Block Recall, Mazes Memory, Listening Recall, Counting Recall, and Backward Digit Recall. Each subtest generates a span score and a trials correct score. From these, standard scores are derived for the three component areas.

The *central executive* is assessed with three dual-task items. For *Listening Recall*, the child determines if a sentence is true or false and then recalls the last word in a series of sentences. For *Counting Recall*, the child counts a set of 4, 5, 6, or 7 dots on a card and then recalls the number of counted dots at the end of a series of cards. *Backward Digit Recall* is a standard format backward digit span.

The *phonological loop* comprises *Digit Recall*, *Word List Recall*, and *Nonword List Recall*, which are standard span tasks. The child repeats stimuli spoken by the tester in the same order. In *Word List Matching*, the tester presents a series of words, beginning with two words and adding one word at each level. The same words, but possibly in a different order, are then presented again, and the child determines if the second list is in the same order as the first list.

The *visuospatial sketchpad* comprises two tasks. *Block Recall* is another span task, but the stimuli consist of a board with nine raised blocks in what appears as a random arrangement. The blocks have numbers on one side that only the tester can see. The tester taps a block (or series of blocks), and the child duplicates the tapping in the same order. For *Mazes Memory*, the tester presents a maze with more than one solution and a picture of an identical maze with a path showing one solution. The picture is removed, and the child duplicates the path in the response booklet. At each level, the complexity of the maze increases.

Processing speed—*WJ-III Visual Matching* (Woodcock et al., 2001) asks examinees to locate and circle two identical numbers in rows of six numbers; examinees have 3 min to complete 60 rows and earn credit by correctly circling the matching numbers in each row. As reported by the test developer, internal consistency reliability is .91.

Domain-Specific Measures of Numerosity

We indexed numerosity with Number Line Estimation, which provides an assessment of the approximate magnitude system, and the Number Sets Test, which provides an assessment of precise representation of precise small quantities and the ability to mentally combine these representations (e.g., ●● ■■■ = 5).

Approximate numerical magnitudes—Following Siegler and Booth (2004), *Number Line Estimation* asks students to estimate where numbers fall on a number line. Stimuli are 24–25-cm number lines, displayed across the center of a standard computer screen. Each number line has a start point of 0 and an endpoint of 100 with a target number printed approximately 5 cm above it in a large font (72 pt). Target numbers are 3, 4, 6, 8, 12, 17, 21, 23, 25, 29, 33, 39, 43, 48, 52, 57, 61, 64, 72, 79, 81, 84, 90, and 96. Stimuli are presented in random order for each child. The tester first explains a number line that includes the 0 and 100 endpoints and is marked in increments of 10. After the tester judges that the child recognizes the concept, a number line containing only the endpoints 0 and 100 is presented, and the child points to where 50 should go. A model number line with the endpoints and the location of 50 marked is then presented, and the child compares his/her response to the model. The tester explains how “the number 50 is half of 100, so we put it halfway in between 0 and 100 on the number line.” Next, the tester teaches the child to use the arrow keys to place a red pointer on the line where 50 should fall on the computer screen. Then, the measure is administered. For each item, the tester says, “If this is zero (pointing) and this is 100 (pointing), where would you put N?” There is no time constraint. We used the absolute value of the difference between the correct placement and the child’s placement (i.e., estimation of accuracy), averaged across trials, which the computer automatically calculates. This estimation accuracy score correlates with mathematics achievement (Geary et al., 2007; Siegler & Booth), and as Siegler and Booth showed, the source of improvement in estimation accuracy is increasing linearity of estimates. Cronbach’s alpha on this sample was .91.

Precise representation of small quantities—The *Number Sets Test* (Geary et al., 2009) assesses the speed and accuracy with which children understand and manipulate small, exact numerosities, while handling those quantities and their Arabic code. Two types of number-set stimuli are used: objects (circles, squares, diamonds, stars) in a 1/2″ square and an Arabic numeral (18 pt font) in a 1/2″ square. Stimuli are joined in domino-like rectangles. The two sides of the “domino” are Objects/Same Objects or Objects/Different Objects or Arabic numerals/Arabic numerals or Objects/Arabic numerals. These dominos are presented in lines of 5 across a page. The last two lines of the page show three 3-square dominos. Target sums (5 or 9) are shown in larger font (36 pt) at the top the page; 5 and 9 were chosen because they represent smaller and larger values within the range of basic Arabic numerals (i.e., 1 to 9). The first page for each target sum contains Objects/Same Objects and Objects/Different Objects; the second page, Objects/Arabic Numerals and Arabic Numerals/Arabic Numerals. On each page, 18 items match the target; 12 are larger than the target; 6 are smaller than the target; and 6 contain 0 or an empty square. The tester begins by explaining two items matching a target sum of 4; then, uses the target sum of 3 for practice, with four lines of two dominos, one is a match and one is not. The measure is then administered. The child is told to move across each line of the page from left to right without skipping any; to “circle any groups that can be put together to make the top number, 5 (9)”; and to “work as fast as you can without making many mistakes.” The child has 60 sec per page for the target sum 5; 90 sec per page for the target sum 9 (time limits chosen to avoid ceiling effects and to assess fluent recognition of quantities). Signal detection methods are applied to the number of hits and false alarms to generate a d' variable that represents the child’s sensitivity to quantities (Geary et al.; MacMillan, 2002). Children who correctly identify many target quantities and commit few false alarms have high d' scores. Children who have as many hits as false alarms have low d' scores, in which case the high number of correct items is due to the child’s bias to respond, not sensitivity to quantity.

Mathematics Outcomes

Number combinations (NCs)—*Math Fact Fluency* (Fuchs, Hamlett, & Powell, 2003) incorporates two subtests. *Addition Fact Fluency* comprises 25 addition fact problems with sums from 5 to 12 (two items have an addend of 1; one item has an addend of zero). *Subtraction Fact Fluency* comprises 25 subtraction fact problems with subtrahends from 5 to 12 (one item has a minuend of 1; one item has a minuend of zero). For each subtest, students have 1 min to write the answers. The score is the number of correct answers across both subtests. Percentage of agreement, calculated on 20% of protocols by two independent scorers, was 97.3. Coefficient alpha on this sample was .94.

Word problems (WPs)—Following Jordan and Hanich (2000), *Word Problems* comprises 14 brief WPs involving change, combine, compare, and equalize relationships. The NCs embedded in the WPs are sums of 7, 8, or 9 or subtrahends of 6, 7, 8, or 9; there are no addends or minuends of zero or one; and answers to the subtraction problems range from 2 to 6. The tester reads each item aloud; students have 30 sec to respond and can ask for re-reading(s) as needed. The score is the number of correct answers. A second scorer independently re-scored 20% of protocols, with agreement of 99.7%. Coefficient alpha on this sample was .84.

Procedure

We assessed students in the fall of first grade. In September, whole-class screening lasted 15 min; individual screening, 5 min. In September and October, we individually tested children on the remaining measures in three 45- to 60-min sessions. Tests were administered by trained examiners, each of whom had demonstrated 100% accuracy in following standard test procedures during mock administrations. All individual sessions were audiotaped, and 15% of tapes, distributed equally across testers, were selected randomly for accuracy checks by an independent scorer. Agreement was between 99.5 and 99.9%. In October, teachers also completed the SWAN Rating Scale on each student.

Results

A principal components factor analysis across WDRB Listening Comprehension and WASI Vocabulary yielded one factor; thus, we created a weighted composite language variable. Each of the remaining constructs was assessed with one measure: for concept formation, WJ-III Concept Formation; for nonverbal problem solving, WASI Matrix Reasoning; for attentive behavior, SWAN; for working memory, the phonological loop, and the visuospatial sketchpad, summary scores from the WMTB-C subtests; for processing speed, WJ-III Visual Matching. The numerosity measures were Number Line Estimation and the Number Sets Test. The two mathematics outcomes were NCs and WPs. In Table 1, we show raw and standard score means and *SD*s, along with zero-order correlations.

Predictors of Number Combination (NC) Skill

We contrasted two models in predicting individual differences in NC skill (see Table 2). In Model 1-NC, we tested the effects of the two numerosity measures, entering Number Line Estimation and the Number Sets Test simultaneously into the regression analysis. Each was a significant predictor in the presence of the other. Overall, Number Line Estimation (negative scores represent accurate placements) and the Number Sets Test (i.e., *d′* score) explained 65.6% of the variance in NC skill, $F(2,202) = 192.75$, $p < .001$. Partitioning this variance revealed that 36.9% of the explained variance was shared between Number Line Estimation and the Number Sets Test, 58.5% was unique to Number Sets Test, and 4.6% was unique to Number Line Estimation.

Model 2-NC shows that each numerosity measure (Number Line Estimation and the Number Sets Test) was a significant predictor of NC skill in the presence of the eight domain-general abilities. The only domain-general ability that was a uniquely significant predictor of NC skill was attentive behavior. Overall, the two numerosity measures and the eight domain-general abilities explained 67.7% of the variance in NC skill; the addition of the eight domain-general abilities did not significantly increase R^2: 2.1% increase, F Change $(8,194) = 1.56$, $p = .141$. Partitioning Model 2-NC variance between the numerosity measures and the domain-general abilities revealed that 62.0% of the explained variance was shared, 34.9% was unique to the numerosity measures, and 3.1% was unique to domain-general abilities.

Predictors of Word-Problem (WP) Skill

We followed the same procedure in predicting individual differences in WP skill (see Table 3). In Model 1-WP, each numerosity measure (Number Line Estimation and the Number Sets Test) was a significant predictor in the presence of the other. Together, Number Line Estimation and the Number Sets Test explained 44.4% of the variance in WP skill, $F(2,202) = 80.66$, $p < .001$. Partitioning this variance revealed that 38.7% of the explained variance was shared between Number Line Estimation and the Number Sets Test, 55.2% was unique to *d′*, and 6.1% was unique to Number Line Estimation.

In Model 2-WP, the Number Sets Test but not Number Line Estimation was a significant predictor of WP skill in the presence of the eight domain-general abilities. The domain-general abilities that were uniquely significant predictors of WP skill were the central executive component of working memory and concept formation, with a trend for language ability ($p = .069$). Overall, the numerosity measures and the eight domain-general abilities explained 55.2% of the variance in WP skill, with the addition of the domain-general abilities resulting in a substantial (10.8%) and significant increase in R^2, F Change $(8,194) = 5.82$, $p < .001$. Partitioning Model 2-WP variance between the numerosity measures and the domain-general abilities revealed that 68.3% of the explained variance was shared, 12.1% was unique to the numerosity measures, and 19.6% was unique to domain-general abilities.

Discussion

We assessed the relative contributions of domain-general abilities and domain-specific numerical competencies on individual differences in first graders' NC and WP skill. We focused first on NC skill because the National Mathematics Advisory Panel (2008) and the National Council of Teachers of Mathematics (2006) identified NC fluency as a core arithmetic skill, and because attaining NC fluency is difficult for many children (e.g., Goldman et al., 1988).

We began by comparing the relative contributions of measures representing the two hypothesized number sense systems to the prediction of individual differences in NC skill.

We entered a measure of approximate representation of large quantities (i.e., Number Line Estimation; Siegler's & Booth, 2004) and a measure of exact representation of small quantities (i.e., *d*′ from the Number Sets Test; Geary et al., 2009) simultaneously into regression analysis predicting NC skill. Each was a significant predictor in the presence of the other, and the two numerosity measures together accounted for a large proportion of the variance (65.6%). This suggests that both core systems hypothesized to comprise number sense contribute to the development of NC skill. Even so, most (58.5% of the 65.6%) of this variance was unique to the measure of exact representations of small quantities (i.e., the Number Sets Test), with 36.9% of the explained variance shared between the two numerosity constructs and only 4.6% attributable specifically to approximate representation of large quantities (i.e., Number Line Estimation).

The importance of the approximate system of number in NC skill development is in keeping with findings of some recent studies (Gilmore et al., 2007; Halberda, Mazzocco, & Feigenson, 2008). Yet, our results suggest that fluency in accessing exact representations of small quantities may be more essential. This makes sense given that NC skill and items on the Number Sets Test require quick recognition and combination of small magnitudes. One possibility is that children who have a strong foundation for verbally subitizing and adding the corresponding quantities are at an advantage in learning NC. This might explain Koontz and Berch's (1996) finding that children with mathematics disability have a smaller subitizing range; it may also explain the common finding that children with mathematics disability have difficulty learning NCs (Geary, 1993). At the same time, it is likely that for other types of mathematics performance, approximate representation of large quantities is the more essential system. In that vein, Booth and Siegler (2008) provided data suggesting the potential importance of Number Line Estimation in predicting students' learning of an addition estimation task. The authors did not, however, consider exact representation of small quantities as a competing predictor. In future studies, researchers might simultaneously consider the two hypothesized core systems of number sense when investigating various aspects of mathematics performance, as done in the present study. In any case, our results support Duncan et al.'s (2008) conclusion about the importance of early numerical competence for supporting formal mathematical learning in school.

After examining the effects of the two numerosity measures on NC skill, we contrasted a model that assessed the effects of the numerosity measures in the presence of the eight domain-general abilities. In this much more stringent test, both numerosity measures retained their status as unique predictors. On this basis, we conclude that each of the two hypothesized core systems of number sense directly affects the development of NC skill even when controlling for the variance associated with a host of domain-general abilities. Just as interesting, however, results showed that the addition of the domain-general abilities resulted in only a small, nonsignificant (2.1%) increase in R^2. This was the case even though several of our measures of domain-general abilities (i.e., central executive, phonological loop, and processing speed) involved numerals. As per Landerl et al. (2004), this creates a stringent test of the domain-specific hypothesis due to the potential for confounding domain-general abilities with numerical processing abilities. In a similar way, because we excluded 73 students due to mathematics performance that was inconsistent with their overall risk status, it is possible that results are biased toward a domain-general perspective, creating an even more stringent test of the domain-specific hypothesis.

Yet, our findings lend support for a domain-specific perspective on NC skill, not only indicating that both hypothesized aspects of numerosity are important in NC skill but also that (a) NC skill does not depend on domain-general abilities above and beyond the influence of numerosity and (b) domain-general abilities do not mediate the effects of numerosity on NC skill. This suggests that NC skill grows out of the development of meaningful and interconnected knowledge about numbers that may have their foundation in children's early ability to subitize and estimate magnitude (e.g., Baroody et al., 2009). In the present study, we did not address whether individual differences in numerosity result from innate differences or from the nature of socially-mediated preschool experiences. Regardless of how these individual differences develop, results indicate the importance of providing strong preschool opportunities to help children construct a strong sense of numerosity.

At the same time, however, we also examined how much of the variance in domain-general abilities and numerosity variables is unique versus shared when explaining NC skill. Partitioning the variance revealed that although 34.9% of the explained variance was unique to the numerosity measures (and only 3.1% was unique to domain-general abilities), 62.0% of the explained variance was shared. Finding that a substantial amount of variance is shared indicates the need for additional research on the viability of a combined perspective. Geary et al. (2009) demonstrated that although performance on the Number Sets Test appears to capture fluency in processing basic numerical information, it is influenced by the central executive and general cognitive ability. Whether this represents method variance in the construction of the Number Sets Test or whether working memory and cognitive ability are important for the functional use of domain-specific number representations during mathematics learning remains to be determined.

Our last question centered on whether the relations for NC skill apply to WP performance, which requires but is not limited to NC skill. When Number Line Estimation and the Number Sets Test were simultaneously entered into regression analysis to predict WP performance, each was a significant predictor in the presence of the other, as was the case for NC skill. However, these numerosity variables accounted for less (44.4%) of the explained variance in WP skill compared to NC performance (65.6%). So it was not surprising that when we assessed the effects of the numerosity measures in the presence of the eight domain-general abilities, this more stringent test revealed that only one of these core systems of number sense, exact representation of small quantities (i.e., d' of the Number Sets Test), not approximate representation of large quantities (i.e., Number Line Estimation), was a unique predictor. This indicates that only exact representation of small quantities exerts a direct effect on WP skill whereas the effects of approximate representation of large quantities are mediated by the domain-general abilities.

In fact, the addition of the set of domain-general measures resulted in a substantial (10.8%) and significant increase in explained variance in WP skill. This stands in stark contrast to the finding that domain-general abilities together did not improve the prediction of NC performance beyond what was captured by the numerosity measures. Moreover, partitioning the variance between the numerosity measures and the domain-general abilities revealed that more variance was attributable to domain-general abilities (19.6%) than to the numerosity measures (12.1%). By contrast, in predicting NC skill, the corresponding percentages were 3.1 and 34.0. In line with the prediction of NC skill, however, a hefty portion (68.3%) of the explained variance was shared between numerosity and domain-general abilities, again suggesting additional research on the viability of a combined perspective.

In conducting research to understand which domain-general abilities may underlie WP skill, our findings indicate the importance of three abilities. First, the central executive component of working memory was uniquely predictive of WP performance, even when two numerosity and seven other domain-general abilities were competing for variance. Theoretical frameworks posit that WPs involve construction of a problem model that requires working memory (e.g., Kintsch & Greeno, 1985). Much of the literature, including the present results, provides support for this theoretical framework (e.g., Fuchs et al., 2005; Swanson & Beebe-Frankenberger, 2004; Swanson & Sachse-Lee, 2001), although findings are not consistent (e.g., Fuchs et al., 2006; Fuchs, Fuchs, Stuebing et al., 2008; Vukovic & Siegal, 2009).

A second potentially important domain-general ability in determining WP skill, as reflected in the present study, is concept formation. As with the central executive, concept formation was uniquely predictive of WP performance, even when two numerosity and seven other domain-general abilities were competing for variance. Concept formation involves identifying, categorizing, and determining rules to classify objects and therefore appears to represent general reasoning or problem-solving ability, as is required in WPs. In prior research, concept formation has also emerged as a determinant of WP skill (Fuchs et al., 2006; Fuchs, Fuchs, Stuebing et al., 2008; Swanson & Sachse-Lee, 2001). Additional studies are needed to determine if concept formation predicts ease of learning and transferring WP schemas (e.g., Fuchs, Fuchs, Craddock et al., 2008; Sweller, 2004).

The present results also suggest the potential importance of a third domain-general ability: language. Although it only approached statistical significance ($p = .069$), we nonetheless discuss this finding because prior studies strongly support the role of language in WP performance, thereby suggesting the appropriateness of a 1-tailed test of significance. Jordan et al. (1995) documented the importance of language when kindergarten and first-grade language-impaired children performed lower than unimpaired peers on WPs. Fuchs et al. (2006) showed its relation to WP skill among third graders while controlling for a host of other variables, and profile analysis (Fuchs, Fuchs, Stuebing et al., 2008) demonstrated that poor language ability was the domain-general ability that reliably distinguished third graders with WP deficits from those with specific calculation deficits. Moreover, it makes sense that language supports WP performance given the obvious need to process linguistic information when building representations of WPs.

In a more general way, the pattern of findings in the present study, whereby the role of numerosity and domain-general abilities differed across NC and WP skill, also suggests that not all aspects of mathematics performance tap the same underlying abilities. These findings lend support for the notion that NC and WP skill may represent distinct domains of mathematical cognition. Results echo Fuchs, Fuchs, Stuebing et al. (2008), in which third-grade students at the lower end of the performance range in calculations versus WPs showed distinctive cognitive profiles, and Hart, Petrill, Thompson, and Plomin (2009), who found that mathematics problem solving has different genetic and environmental influences than mathematics calculation. Together, these studies suggest that the development of mathematical cognition does not depend on a uniform set of numerical or general cognitive competencies but rather that the relative contributions of numerosity and domain-general explanations to the development of mathematical cognition differ depending on the nature of mathematics performance. In fact, elementary school mathematics is conceptualized in strands (e.g., concepts, numeration, measurement, arithmetic, algorithmic computation, and problem solving), and high-school curriculum offerings include algebra, geometry, trigonometry, and calculus. Yet, little is understood about how aspects of mathematical cognition relate to each other: which are shared or distinct or how development in one domain corresponds with development in another or how learning in one domain promotes learning in another. Research is needed to expand theoretical insight into the nature of these types of mathematical competence and their development.

Acknowledgments

This research was supported in part by Award Number RO1 HD053714 and by Core Grant #HD15052 from the National Institute of Child Health and Human Development (NICHD) to Vanderbilt University and Award Number R37 HD045914 co-funded by NICHD and the Office of Special Education and Rehabilitation Services award to the University of Missouri. The content is solely the responsibility of the authors and does not necessarily represent the

official views of the National Institute of Child Health and Human Development or the National Institutes of Health.

References

Ackerman PL, Beier ME, Boyle MO. Working memory and intelligence: The same or different constructs? Psychological Bulletin. 2005; 131:30–60. [PubMed: 15631550]

Baddeley, AD. Working memory. Oxford: Oxford University Press; 1986.

Baroody AJ. Mastery of basic number combinations: Internalization of relationships or facts? Journal of Research on Mathematics Education. 1985; 16:83–98.

Baroody AJ, Bajwa NP, Eiland M. Why can't Johnny remember the basic facts? Developmental Disabilities Research Reviews. 2009; 15:69–79. [PubMed: 19213010]

Baroody, AJ.; Lai, M-L.; Mix, KS. The development of number sense and operation sense in early childhood. In: Saracho, O.; Spodek, B., editors. Handbook of research on the education of young children. Mahwah, NJ: Erlbaum; 2006. p. 187-221.

Barrouillet P, Fayol M, Lathuliére E. Selecting between competitors in multiplication tasks: An explanation of the errors produced by adolescents with learning disabilities. International Journal of Behavioral Development. 1997; 21:253–275.

Booth JL, Siegler RS. Numerical magnitude representations influence arithmetic learning. Child Development. 2008; 79:1016–1031. [PubMed: 18717904]

Bull R, Johnston RS. Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. Journal of Experimental Child Psychology. 1997; 65:1–24. [PubMed: 9126630]

Cantlon JF, Platt ML, Brannon EM. Beyond the number domain. Trends in Cognitive Science. 2009; 13:83–91.

Case, R. Intellectual development: Birth to adulthood. San Diego: Academic; 1985.

Connolly, AJ. KeyMath-Revised. Circle Pines, MN: American Guidance; 1998.

Domahs F, Delazier M. Some assumptions and facts about arithmetic facts. Psychology Science. 2005; 47:96–111.

Duncan GJ, Dowsett CJ, Classens A, et al. School readiness and later achievement. Developmental Psychology. 2008; 43:1428–1446. [PubMed: 18020822]

Embretson SE. The role of working memory capacity and general control processes in intelligence. Intelligence. 1995; 20:169–189.

Engle RW, Tuholski SW, Laughlin JE, Conway ARA. Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. Journal of Experimental Psychology: General. 1999; 128:309–331. [PubMed: 10513398]

Feigenson L, Dehaene S, Spelke E. Core systems of number. TRENDS in Cognitive. Science. 2004; 8:307–314.

Fuchs D, Fuchs LS, Thompson A, Al Otaiba S, Yen L, Yang N, Braun M, O'Connor R. Is reading important in reading-readiness programs?: A randomized field trial with teachers as program implementers. Journal of Educational Psychology. 2001; 93:251–267.

Fuchs LS, Compton DL, Fuchs D, Paulsen K, Bryant JD, Hamlett CL. The prevention, identification, and cognitive determinants of math difficulty. Journal of Educational Psychology. 2005; 97:493–513.

Fuchs LS, Fuchs D, Compton DL. Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. Exceptional Children. 2004; 71:7–21.

Fuchs LS, Fuchs D, Compton DL, Powell SR, Seethaler PM, Capizzi AM, Schatschneider C, Fletcher JM. The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. Journal of Educational Psychology. 2006; 98:29–43.

Fuchs LS, Fuchs D, Craddock C, Hollenbeck KN, Hamlett CL, Schatschneider C. Effects of small-group tutoring with and without validated classroom instruction on at-risk students' math problem solving: Are two tiers of prevention better than one? Journal of Educational Psychology. 2008; 100:491–509. [PubMed: 19122881]

Fuchs LS, Fuchs D, Stuebing K, Fletcher JM, Hamlett CL, Lambert WE. Problem-solving and computation skills: Are they shared or distinct aspects of mathematical cognition? Journal of Educational Psychology. 2008; 100:30–47. [PubMed: 20057912]

Fuchs, LS.; Hamlett, CL.; Fuchs, D. First Grade Test of Computational Fluency and First-Grade Test of Mathematics Concepts and Applications. L.S. Fuchs; 228 Peabody, Vanderbilt University, Nashville, TN 37203: 1990.

Fuchs, LS.; Hamlett, CL.; Powell, SR. Math Fact Fluency. L.S. Fuchs; 228 Peabody, Vanderbilt University, Nashville, TN 37203: 2003.

Gallistel CR, Gelman R. Preverbal and verbal counting and computation. Cognition. 1992; 44:43–74. [PubMed: 1511586]

Geary DC. Mathematical disabilities: Cognitive, neuropsychological, and genetic components. Psychological Bulletin. 1993; 114:345–362. [PubMed: 8416036]

Geary DC. An evolutionary perspective on learning disability in mathematics. Developmental Neuropsychology. 2007; 32:471–519. [PubMed: 17650991]

Geary DC, Bailey DH, Hoard MK. Predicting mathematical achievement and mathematical learning disability with a simple screening tool: The Number Sets Test. Journal of Psychoeducational Assessment. 2009; 27:XX.

Geary DC, Hoard MK, Byrd-Craven J, Nugent L, Numtee C. Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. Child Development. 2007; 78:1343–1459. [PubMed: 17650142]

Geary DC, Lin J. Numerical cognition: Age-related differences in the speed of executing biologically-primary and biologically-secondary processes. Experimental Aging Research. 1998; 24:101–137. [PubMed: 9555566]

Gilmore CK, McCarthy SE, Spelke ES. Symbolic arithmetic knowledge without instruction. Nature. 2007 May 31.447:589–591. [PubMed: 17538620]

Goldman SR, Pellegrino JW, Mertz DL. Extended practice of addition facts: Strategy changes in learning-disabled students. Cognition and Instruction. 1988; 5:223–265.

Halberda J, Mazzocco M, Feigenson L. Individual differences in non-verbal number acuity correlate with maths achievement. Nature. 2008 October 2.455:665–668. [PubMed: 18776888]

Hanich LB, Jordan NC, Kaplan D, Dick J. Performance across different areas of mathematical cognition in children with learning disabilities. Journal of Educational Psychology. 2001; 93:615–626.

Hart SA, Petrill AS, Thompson LA, Plomin R. The ABCs of math: A genetic analysis of mathematics and its links with reading ability and general cognitive ability. Journal of Educational Psychology. 2009; 101:388–402. [PubMed: 20157630]

Hecht SA, Torgesen JK, Wagner RK, Rashotte CA. The relations between phonological processing abilities and emerging individual differences in mathematical computational skills: A longitudinal study from second to fifth grades. Journal of Experimental Child Psychology. 2001; 79:192–227. [PubMed: 11343408]

Hitch GJ, McAuley E. Working memory in children with specific arithmetical learning disabilities. British Journal of Psychology. 1991; 82:375–386. [PubMed: 1954527]

Jordan NC, Hanich L. Mathematical thinking in second-grade children with different forms of LD. Journal of Learning Disabilities. 2000; 33:567–578. [PubMed: 15495398]

Jordan NC, Levine SC, Huttenlocher J. Calculation abilities in young children with different patterns of cognitive functioning. Journal of Learning Disabilities. 1995; 28:53–64. [PubMed: 7844488]

Kail R. Developmental change in speed of processing during childhood and adolescence. Psychological Bulletin. 1991; 109:490–501. [PubMed: 2062981]

Kintsch W, Greeno JG. Understanding and solving word arithmetic problems. Psychological Review. 1985; 92:109–129. [PubMed: 3983303]

Koontz KL, Berch DB. Identifying simple numerical stimuli: Processing inefficiencies exhibited by arithmetic learning disabled children. Mathematical Cognition. 1996; 2:1–23.

Landerl K, Bevan A, Butterworth B. Developmental dyscalculia and basic numerical capacities: A study of 8–9 year old students. Cognition. 2004; 93:99–125. [PubMed: 15147931]

Logie RH, Baddeley AD. Cognitive processes in counting. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1987; 13:310–326.

Mandler G, Shebo BJ. Subitizing: An analysis of its component processes. Journal of Experimental Psychology General. 1982; 111:1–21. [PubMed: 6460833]

MacMillan, NA. Signal detection theory. In: Wixted, J.; Pashler, H., editors. Stevens' handbook of experimental psychology: Methodology in experimental psychology. 3. Vol. 4. New York: John Wiley & Sons; 2002. p. 43-90.

Mix, KS.; Huttenlocher, J.; Levine, SC. Math without words: Quantitative development in infancy and early childhood. New York: Oxford; 2002.

National Council of Teachers of Mathematics. Curriculum focal points for prekindergarten through grade 8 mathematics. Reston, VA: National Council of Teachers of Mathematics; 2006. Retrieved January 9, 2009 from http://www.nctm.org/focalpoints/downloads.asp

National Mathematics Advisory Panel. Foundations for Success: Final Report of the National Mathematics Advisory Panel. Washington, DC: United States Department of Education; 2008. http://www.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf

Pickering, S.; Gathercole, S. Working Memory Test Battery for Children. London: The Psychological Corporation; 2001.

Siegler RS, Booth JL. Development of numerical estimation in young children. Child Development. 2004; 75:428–444. [PubMed: 15056197]

Siegler RS, Opfer J. The development of numerical estimation: Evidence for multiple representations of numerical quantity. Psychological Science. 2003; 14:237–243. [PubMed: 12741747]

Siegler RS, Ramani GB. Playing board games promotes low-income children's numerical development. Developmental Science. 2008; 11:655–661. [PubMed: 18801120]

Swanson HL, Beebe-Frankenberger M. The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious mathematics difficulties. Journal of Educational Psychology. 2004; 96:471–491.

Swanson HL, Sachse-Lee C. Mathematical problem solving and working memory in children with learning disabilities: Both executive and phonological processes are important. Journal of Experimental Child Psychology. 2001; 79:294–321. [PubMed: 11394931]

Swanson, J., et al. Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and the SWAN rating scales. 2004. Downloaded from www.adhd.net on 12/20/2004

Sweller J. Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. Instructional Science. 2004; 32:9–31.

The Psychological Corporation. Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: Harcourt Brace & Company; 1999.

vanMarle K, Wynn K. Infants' auditory enumeration: Evidence for analog magnitudes in the small number range. Cognition. 2009; 111:302–316. [PubMed: 19324330]

Walberg HJ. Improving the productivity of America's schools. Educational Leadership. 1984; 41:19–27.

Wechsler, D. Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: Psychological Corporation; 1999.

Wilkinson, GS. Wide Range Achievement Test 3. Wilmington: Wide Range; 1993.

Woodcock, RW. Woodcock Diagnostic Reading Battery. Itasca, IL: Riverside; 1997.

Woodcock, RW. Woodcock Reading Mastery Test - Revised. Circle Pines, MN: American Guidance Service; 1998.

Woodcock, RW.; McGrew, KS.; Mather, N. Woodcock-Johnson III. Itasca, IL: Riverside; 2001.

Xu F. Numerosity discrimination in infants: Evidence for two systems of representation. Cognition. 2003; 89:B15–B25. [PubMed: 12893126]

Zhu, J. WASI manual. San Antonio: The Psychological Corporation; 1999.

Table 1

Means, Standard Deviations, and Correlations among Domain-General Ability, Numerosity, and Mathematics Outcome Variables (n-205)

	Raw Score		Standard Score		Correlations													
					Language													
Variables	**Mean**	**(*SD*)**	**Mean**	**(*SD*)**	**LF**	**LC**	**V**	**CF**	**NP**	**IA**	**EF**	**PL**	**VS**	**PS**	**NS**	**NL**	**NCs**	**WP**
Domain-General Abilities																		
Language Factor (LF)	0.01	(1.01)	--		-													
WDRB Listening Comprehension (LC)	14.00	(5.16)	87.18	(16.34)	88	-												
WASI Vocabulary (V)	17.25	(6.65)	41.22	(10.50)	87	61	-											
Concept Formation (CF)	9.42	(5.47)	95.00	(13.44)	52	54	44	-										
Nonverbal Problem Solving (NP)	8.70	(5.43)	49.01	(9.57)	42	36	42	50	-									
Attentive Behavior (IA)	38.99	(12.65)	--		49	46	45	47	42	-								
WMTB Executive Function (EF)	--		76.24	(18.28)	55	50	42	48	51	48	-							
WMTB Phonological Loop (PL)	--		96.60	(17.65)	43	39	25	33	27	31	49	-						
WMTB Visuospatial Sketchpad (VS)	--		78.56	(15.14)	30	24	21	31	42	32	47	32	-					
Processing Speed (PS)	474.37	(7.28)	95.80	(15.38)	36	36	35	43	43	44	44	24	37	-				
Numerosity Measures																		
Number Sets (NS)	0.05	(1.21)	--		47	45	47	59	51	59	53	28	35	51	-			
Number Line Estimation (NL)	22.48	(7.75)	--		45	37	51	34	40	39	37	14	27	36	47	-		
Mathematics Outcomes																		
Number Combinations (NCs)	20.44	(12.15)	--		48	43	49	54	46	58	47	30	35	50	79	52	-	
Word Problems (WP)	3.83	(3.04)	--		54	49	51	58	49	52	59	31	34	42	65	45	66	

Standard scores are mean=100 (*SD*=15) except for Wechsler Abbreviated Intelligence Scale Vocabulary and nonverbal problem solving, where mean=50 (*SD*=10). Multiply correlations by .01. Correlations involving Number Line Estimation are negative. WASI is Wechsler Abbreviated Intelligence Scale. WDRB is Woodcock Diagnostic Reading Battery. WJ III is Woodcock-Johnson Psycho-Educational Battery–Revised. Language is WDRB Listening Comprehension and WASI Vocabulary; Concept Formation is WJ III Concept Formation; Nonverbal problem solving is WASI Matrix Reasoning; WMTB is Working Memory Test Battery – Children; Attentive behavior is the SWAN; Processing Speed is WJ III Visual Matching; Arithmetic is the Test of Mathematics Fact Fluency.

Table 2

Regression Models Predicting Individual Differences in First-Grade Number Combination Skill

Model	R^2	SEE	R^2 Change	*F* Change	*p*-value	B	SE	Beta	*t*	p-value
1: Numerosity	.656	7.16	.656	(2,202) 192.75	< .001					
Constant						27.03	1.73		15.66	< .001
Number Line Estimation						−0.31	0.07	−.20	− 4.21	< .001
Number Sets Test						7.06	0.47	.70	15.02	< .001
2. Numerosity + Domain-General	.677	7.08	.021	(8,194) 1.56	.141					
Constant						14.64	6.03		2.43	.016
Language						0.43	0.69	.04	0.62	.535
Attentive Behavior						0.11	0.05	.12	2.18	.031
Central Executive						−0.41	0.04	−.06	−1.03	.305
Phonological Loop						0.04	0.03	.05	1.06	.290
Visuospatial Sketchpad						0.03	0.04	.04	0.74	.459
Concept Formation						0.03	0.05	.03	0.53	.594
Nonverbal Reasoning						−0.03	0.07	−.02	−0.37	.712
Processing Speed						0.03	0.04	.04	0.85	.394
Number Line Estimation						−0.27	0.08	−.17	−3.46	.001
Number Sets Test						6.11	0.59	.61	10.37	< .001

Table 3

Regression Models Predicting Individual Differences in First-Grade Word-Problem Skill

Model	R^2	SEE	R^2 Change	*F* Change	*p*-value	B	SE	Beta	*t*	p-value
1: Numerosity	.444	2.28	.444	(2,202) 80.66	< .001					
Constant						5.47	0.54		10.20	< .001
Number Line Estimation						−0.76	0.23	−.20	− 3.37	.001
Number Sets Test						1.39	0.15	.55	9.49	< .001
2. Numerosity + Domain-General	.552	2.09	.108	(8,194) 5.82	< .001					
Constant						−1.04	1.78		− 0.58	.560
Language						0.37	0.20	.12	1.83	.069
Attentive Behavior						0.02	0.02	.08	1.26	.209
Central Executive						0.04	0.01	.23	3.30	.001
Phonological Loop						−0.01	0.01	−.03	− 0.56	.578
Visuospatial Sketchpad						−0.00	0.01	−.01	− 0.09	.929
Concept Formation						0.03	0.01	.14	2.24	.026
Nonverbal Reasoning						0.01	0.02	.02	0.34	.736
Processing Speed						−0.01	0.01	−.04	− 0.73	.467
Number Line Estimation						−0.04	0.02	−.10	− 1.66	.100
Number Sets Test						0.80	0.17	.32	4.63	< .001