Journal of Experimental Psychology: General

Stability and Change in Markers of Core Numerical Competencies

Robert Reeve, Fiona Reynolds, Judi Humberstone, and Brian Butterworth

Online First Publication, March 12, 2012. doi: 10.1037/a0027520

Journal of Experimental Psychology: General

0096-3445/12/$12.00 DOI: 10.1037/a0027520

Stability and Change in Markers of Core Numerical Competencies

Robert Reeve, Fiona Reynolds, and Judi Humberstone
University of Melbourne

Brian Butterworth
University of Melbourne and University College London

Dot enumeration (DE) and number comparison (NC) abilities are considered markers of core number competence. Differences in DE/NC reaction time (RT) signatures are thought to distinguish between typical and atypical number development. Whether a child's DE and NC signatures change or remain stable over time, relative to other developmental signatures, is unknown. To investigate these issues, the DE and NC RT signatures of 159 children were assessed 7 times over 6 years. Cluster analyses identified within-task and across-age subgroups. DE signatures comprised 4 parameters: (a) the RT slope within the subitizing range, (b) the RT slope for the counting range, (c) the subitizing range (indicated by the point of slope discontinuity), and (d) the overall average DE RT response. NC RT signatures comprised 2 parameters (NC intercept and slope) derived from RTs comparing numbers 1 to 9. Analyses yielded 3 distinct DE and NC profiles at each age. Within-age subgroup profiles reflected differences in 3 of the 4 DE parameters and only 1 NC parameter. Systematic changes in parameters were observed across ages for both tasks, and both tasks broadly identified the same subgroups. Sixty-nine percent of children were assigned to the same subgroup across age, even though parameters themselves changed across age. Subgroups did not differ in processing speed or nonverbal reasoning, suggesting that DE and NC do not tap general cognitive abilities but reflect individual differences specific to the domain of numbers. Indeed, both DE and NC subgroup membership at 6 years predicted computation ability at 6 years, 9.5 years, and 10 years.

Keywords: core number, dot enumeration, number comparison

Supplemental materials: http://dx.doi.org/10.1037/a0027520.supp

The abilities to identify, order, and compare numerical quantities are core aspects of number competence (Berch, 2005; Butterworth, 1999, 2005a, 2005b, 2010; Dehaene, 1997; Desoete, Roeyers, & De Clercq, 2004; Fuson, 1998; Gersten, Jordan, & Flojo, 2005; Laski & Siegler, 2007). Object enumeration, typically visual arrays of dots, and number comparison (NC) tasks are used to study these abilities. NC studies assess the speed and accuracy with which the relative magnitude of two numerical values is identified (e.g., "which number is larger"). Numbers that are closer in magnitude are judged more slowly than those that are more distant in magnitude (often referred to as the *symbolic distance effect*; Moyer & Landauer, 1967). It has been found that performance on NC, using either or both number symbols (e.g., digits) or arrays of dots are associated with arithmetic competence (Holloway & Ansari, 2009; Mazzocco, Feigenson, & Halberda, 2011). For example, Halberda, Mazzocco, and Feigenson (2008) have found that accuracy in comparing two arrays of large numbers of dots correlates with measures of school arithmetic; and Piazza et al. (2010) found that children identified as dyscalculic, a congenital difficulty in learning arithmetic, were significantly less accurate than age-matched controls. Koontz and Berch (1996) and Landerl, Bevan, and Butterworth (2004) have found that a disability in reporting accurately the number of dots in an array is a marker of dyscalculia. These studies suggest that both NC and enumeration require access to representations of numerical magnitude that form the basis of arithmetic. Digit comparison involves accessing two magnitude representations from the digit symbols, whereas dot enumeration (DE) involves matching a magnitude representation to a word symbol. Both matching tasks thus involve a common core representation of numerical magnitude. In this article we consider symbolic number judgments and enumeration as indices of core number abilities.

Object enumeration and NC can thus be used as simple markers of the capacity to acquire arithmetic that are relatively independent of school experiences. These markers have the potential to be diagnostically valuable in identifying learners who are likely to have trouble with arithmetic in school. However, for this to be the case, it is necessary to establish whether they are stable over time for individual children relative to other children.

Items on standardized tests of math ability tend to reflect normative, age-related performance criteria and ipso facto differ across age (Butterworth, 2005a). They also frequently depend on knowledge of formal mathematical procedures. Even easy items may involve formal calculation skills. Similarly, performance on some well-studied indices of math competence (e.g., single-digit addition) requires practice, and for many children, ceiling perfor-

Robert Reeve, Fiona Reynolds, and Judi Humberstone, School of Psychological Sciences, University of Melbourne, Melbourne, Australia; Brian Butterworth, School of Psychological Sciences, University of Melbourne, Melbourne, Australia, and Institute of Cognitive Neuroscience, University College London, London, England.

The research reported herein was supported by a grant to Robert Reeve and Brian Butterworth from the Australian Research Council.

Correspondence concerning this article should be addressed to Robert Reeve, Melbourne School of Psychological Sciences, University of Melbourne VIC 3010, Australia. E-mail: r.reeve@unimelb.edu.au

mance is achieved relatively young. This is not to say that the acquisition process itself is uninteresting. However, both standardized and single-digit tests depend on formal schooling, and typical and atypical performances likely reflect a complex mix of experiential and cognitive factors. Disentangling the impact of these factors on children's numerical abilities has proved challenging, and identifying age-independent indices of math abilities would provide a framework for investigating the development of mathematical competence.

To evaluate the theoretical and practical value of a marker of a cognitive capacity, it is important to track its developmental trajectory (Ansari, 2010; Ansari & Karmiloff-Smith, 2002). Although NC and object enumeration are often included in screening tests designed to identify math learning difficulties in children (Butterworth, 2005a, 2005b; Chard et al., 2005; Desoete, Ceulemans, Roeyers, & Huylebroeck, 2009; Stock, Desoete, & Roeyers, 2010), little work has investigated the relative stability or change over time of within-task parameters (e.g., reaction times [RTs]) in these two tasks. Indeed, most previous longitudinal work has reported global measures of performance (usually RT or correct performances) and has paid little attention to the stability or otherwise of within-task parameters. Here we report patterns of stability and change in both abilities in a large sample of children over a 6-year period. Our aim is to explore the consistency of within-task indices associated with subitizing and NC performances.

Dot Enumeration and Subitizing

In DE tasks, the aim is to assess the speed and accuracy with which different numbers of dots are identified. Small numbers of dots ($n \leq 4$) are typically enumerated rapidly and accurately, whereas larger numbers of dots ($n \geq 4$) are enumerated more slowly and less accurately. These behavioral differences, along with associated neurological data, are regarded as evidence for two distinct enumeration systems (Vetter, Butterworth, & Bahrami, 2011). The ability to rapidly and accurately identify small numbers of dots is described as *subitizing* and is distinct from the slower sequential process of counting (Butterworth, 1999; Schleifer & Landerl, 2011). Surprisingly, little attention has been paid to individual differences in subitizing or how subitizing abilities change over time. Characterizing the developmental nature of subitizing may provide valuable information for assessing the cognitive basis of subitizing abilities.

Four DE RT parameters may be important: (a) the RT slope within the subitizing range, (b) the RT slope for the counting range, (c) the subitizing range (indicated by the point of slope discontinuity), and (d) the overall average DE RT response. For items in the subitizing range (typically one to four) the slope function is typically shallow (increments of 40–120 ms per item; in the counting range the increment tends to be 250–350 ms per item; Trick & Pylyshyn, 1993). Surprisingly, researchers have rarely investigated the nature of changes in these parameters with age. Trick and Pylyshyn (1994) speculated that children's growing familiarity with the number system should aid the association of number names with discrete quantities and lead to flatter slopes in the subitizing range. Trick, Enns, and Brodeur (1996) and Basak and Verhaeghen (2003) found that with age, children displayed shallower slopes in the subitizing range but not the counting range. Currently, it is unclear how the four DE parameters are related.

The apparent failure to subitize small numerosities (counting them instead) has been implicated in several cognitive disorders and is associated with dyscalculia. Subitizing deficits have mostly been associated with right parietal disruptions, particularly the intraparietal sulcus. These deficits have been shown in adults with Turner's syndrome (Bruandet, Molko, Cohen, & Dehaene, 2004), children with cerebral palsy (Arp & Fagard, 2005; Arp, Taranne, & Fagard, 2006), velocardiofacial syndrome (a.k.a., chromosome 22q11.2 deletion syndrome, or DS22q11.2; De Smedt et al., 2007; Simon, Bearden, McDonald-McGinn, & Zackai, 2005; Simon et al., 2008), fragile X syndrome, and Williams syndrome (Mazzocco & Hanich, 2010; Paterson, Girelli, Butterworth, & Karmiloff-Smith, 2006). Subitizing impairments have also been observed in adult individuals with acquired Gerstmann's syndrome (Cipolotti, Butterworth, & Denes, 1991; Lemer, Dehaene, Spelke, & Cohen, 2003), who appear to count individual items in arrays of less than four and are poor calculators. Similarly, children who show a constant linear RT increase with no point of discontinuity when enumerating successive numerosities (i.e., who are not subitizing small numerosities) are also very poor at arithmetic (Arp & Fagard, 2005; Arp et al., 2006; Koontz & Berch, 1996; Landerl, Bevan, & Butterworth, 2004).

The most obvious feature of the RT function for DE is the change in enumeration latency as a function of array size, a discontinuity claimed to represent the boundary between subitizing and counting (Trick & Pylyshyn, 1993). A recent neuroimaging study has identified activations specific for enumerations in the subitizing range in an area in the right temporoparietal junction (Vetter et al., 2011). The point of discontinuity is claimed to be around either three (Logan & Zbrodoff, 2003; Trick et al., 1996; Watson, Maylor, & Bruce, 2005; Wender & Rothkegel, 2000) or four (Piazza, Mechelli, Butterworth, & Price, 2002). It is usually determined statistically as the point at which successive RT increases in the enumeration of numerosities change from a linear to a nonlinear function (Piazza, Giacomini, Le Bihan, & Dehaene, 2003; Simon et al., 2005; Svenson & Sjoberg, 1983; Trick & Pylyshyn, 1994; Tuholski, Engle, & Baylis, 2001). Svenson and Sjoberg (1983) examined the subitizing ranges and differences in enumeration processes of children and adults (7-year-olds to adults) and found that, in general, children subitized to three dots and older children and adults subitized to four, although there was significant within-age variability. The RT enumeration slope and intercept for children within the subitizing range decreased with age, with the greatest decrement occurring at 8 years of age; however, little change occurred after 10 years of age.

Number Comparison

NC tasks have long been used to examine the nature of magnitude representations and the emergence of visuospatial representations of numbers. They comprise both nonsymbolic and symbolic versions, both of which are considered to be represented in terms of an ordered, approximate number system (i.e., the mental number line; Mazzocco et al., 2011). Although the former is thought to reflect magnitude representation per se, the latter is claimed to depend on the putative ability to link Arabic symbols to magnitude representations. Although difficulties making symbolic NC judgment may reflect symbolic access difficulties, children's familiarity with single digit Arabic number is likely to minimize this effect. Indeed, differences in both nonsymbolic (e.g., Mazzocco et al., 2011) and symbolic (e.g., Landerl et al., 2004) NC judgment tasks have been shown to predict differences in arithmetic skills.

Typically, symbolic number tasks have been included in math screeners for children on the assumption that NC ability is an important index of math ability (e.g., Butterworth, 2003). NC tasks are characterized by the slope function and intercept of response time according to the numerical distance between number pairs. Developmental studies show that the intercept and linear slope of numerical difference judgments decreases from kindergarten to fourth grade, with minor changes occurring thereafter (Duncan & McFarland, 1980; Sekuler & Mierkiewicz, 1977).

NC abilities have been studied developmentally (Baker et al., 2002; Chard et al., 2005; Clark & Shinn, 2004; Geary, Hamson, & Hoard, 2000; Geary, Hoard, & Hamson, 1999; Gersten et al., 2005; Jordan, Kaplan, Nabors Olah, & Locuniak, 2006; Locuniak & Jordan, 2008; Mazzocco & Thompson, 2005; Shalev, Manor, Auerbach, & Gross-Tsur, 1998; Shalev, Manor, & Gross-Tsur, 1997, 2005; Silver, Pennett, Black, Fair, & Balise, 1999). Different number distance effects (different NC slopes) have been observed in children with math learning deficits (Mussolin, Mejias, & Noel, 2010; Rousselle & Noël, 2007) and comorbid visuospatial deficits (Bachot, Gevers, Fias, & Roeyers, 2005). Deficits in accuracy have only been observed for close large number pairs (Geary et al., 2000; Rousselle & Noël, 2007), consistent with Sekuler & Mierkiewicz's (1997) suggestion that the analogue representation of numbers is either compressed and/or displays greater dispersion around numerical values in early development (Dehaene, 2003; Gallistel & Gelman, 1992). Faster and more accurate judgments of numerical magnitude may thus reflect the growing understanding of cardinal relationships, improvements in transcoding, and automaticity in accessing numerical information (Girelli, Lucangelli, & Butterworth, 2000; Rubinsten, Henik, Berger, & Shahar-Slavev, 2002). However, patterns of development of slope and intercept in NC have not been investigated longitudinally. This would help determine whether differences between individuals are stable over time or reflect age-related changes.

Factors Affecting Dot Enumeration and Number Comparison Abilities

General cognitive competencies (e.g., processing speed, executive function abilities) have been found to be associated with children's math learning abilities: *processing speed* (Kail & Ferrer, 2007; Salthouse & Davis, 2006), *executive functions* (Bull, Johnson, & Roy, 1999; Geary & Brown, 1991; Geary & Lin, 1996). However, findings have been inconsistent, and not all studies report a link between math ability and general cognitive competence (Butterworth, 2005a). Nevertheless, improvement in processing speed and working memory may be lead to increases in subitizing span, but few studies have investigated this possibility (however, see Green & Bavelier, 2006; Tuholski et al., 2001).

Characterizing Different Performance Profiles

Researchers tend to focus on age-related changes in competencies on the assumption that developmental differences in abilities have been identified. Analyses are typically based on aggregate data, with interindividual variability being regarded as noise. This practice has been criticized by some researchers who argue that there might be circumstances where this assumption is unwarranted or misleading (Chochon, Cohen, van de Moortele, & Dehaene, 1999; Dehaene, Piazza, Pinel, & Cohen, 2003; Rueckert et al., 1996; Weinert & Helmke, 1998). Indeed, the meaning of large within-age variability in performance on most math tasks is rarely explicated (Ansari, 2010; Ansari & Karmiloff-Smith, 2002; Rosengren & Braswell, 2001). Variation in RTs in the subitizing range in adults, for example, has been regarded as interindividual noise (Akin & Chase, 1978; Balakrishnan & Ashby, 1991, 1992; Piazza et al., 2003; Trick & Pylyshyn, 1994). However, Balakrishnan and Ashby (1991) conceded that different subitizing profiles may be embedded within the overall putative interindividual noise profile.

Research Focus

We ask whether meaningful subgroups are embedded within an overall response distribution and whether subgroup assignment is consistent over time. We focus on two general questions: First, do children remain in the same subgroup over time, independent of changes that may occur within groups (e.g., speed of responding); and second, even though subgroup assignment may remain constant, do systematic changes occur within and across subgroups over time? Moreover, insofar as DE and NC are core markers of numerical competence, we hypothesize that subgroup membership will remain constant over time. In particular, it has been assumed, but not demonstrated, that DE and NC tap the same representations of numerical magnitude, although not in the same way. In other words, the two tasks measure the same construct using different processes. Therefore, it is important to establish the degree to which both tasks identify the same subgroups over time.

Although the primary analytic focus is on the consistency of DE and NC subgroup membership over time, we also consider the relationship between subgroup membership and general cognitive abilities (basic RT and Ravens Colored Progressive Matrices [RCPM]; Raven, Court, & Raven, 1986). Insofar as the two cognitive markers (basic RT and RCPM) are associated with subgroup membership, it would suggest that DE and NC abilities reflect general cognitive abilities, rather than domain-specific indices of numerical competence. Finally, we wished to determine whether subgroup membership predicts arithmetical computation abilities at 6 years, 9.5 years, and 10 years.

Method

Participants

The sample comprised one hundred fifty-nine 5.5- to 6.5-year-olds (overall M = 72.53 months, SD = 4.55 months): 95 boys (M = 73.15 months, SD = 4.57 months) and 64 girls (M = 71.61 months, SD = 4.57 months). Children attended one of seven independent schools in middle-class suburbs of a large Australian city and, at the beginning of the study, were halfway through their first year of formal schooling. (Approximately 25% of children attend fee-paying independent schools at the primary level in Australia, which increases to 38% at the secondary school level.) The sample was specifically selected to minimize socioeconomic class effects, often associated with cognitive competence (Jordan, Huttenlocher, & Levine, 1992; Sirin, 2005), and to minimize sample attrition. Although children came from different ethnic/cultural backgrounds, all spoke English fluently; had normal or corrected-to-normal vision; and, according to school personnel, had no known learning difficulties. The ethical requirements of the authors' university were followed in conducting the research.

Children were interviewed individually on seven occasions over a 6-year period as part of a larger study. On each occasion they completed a series of tests, including those reported here (see Appendix A in the online supplemental materials for a complete list of tasks used in this longitudinal study). Times between interviews for individual children varied slightly because of the time taken to interview the entire sample (approximately 3 months). The mean ages for the test times were (a) 6 years (5.5–6.5 years) $M = 72.15$ months; (b) 7 years (6.5–7.5 years) $M = 85.51$ months; (c) 8.5 years (8–9 years) $M = 104.95$ months; (d) 9 years (8.5–9.5 years) $M = 110.60$ months; (e) 9.5 years (9–10 years) $M = 116.77$ months; (f) 10 years (9.5–10.5 years) $M = 122.77$ months; and (g) 11 years (10.5–11.5 years) $M = 129.45$ months.

Procedures and Materials

Data for the initial phase of the study were collected in two sessions, approximately 1 week apart. In the first, children were familiarized with our computer-based, RT methods and practiced enumerating canonical dot arrays. The RT familiarity task was always presented first. In the second session, the speed and accuracy with which children completed enumerating random dot arrays and NC judgments were assessed. Both sessions lasted approximately 20 min. The purpose of the familiarization procedures was to introduce the children to the stimuli presentations and to allow them to practice recording their NC judgments by pressing the appropriate computer key.

Speed and accuracy enumerating random dot arrays and making NC judgments were collected for all the remaining phases of the study. It was not considered necessary to run practice tasks again in later test phases. At 6 years, the speed at which children named numbers and letters was measured. At 8 and 9.5 years, children completed a RT task to assess their basic reaction times. At 9.5 years, they completed the Ravens Colored Progressive Matrices test, a measure of nonverbal reasoning ability. At 6 years, children completed single-digit addition; at 9.5 years, they completed a double-digit arithmetic test (addition, subtraction, and multiplication); and at 10 years, they completed a multidigit computation test (three-digit subtraction, multiplication, and division) to assess their arithmetic calculation abilities.

In all tasks in which RT was recorded, stimulus presentation was controlled by a PC running DMDX (Version 2; Forster & Forster, 2001), and stimuli were presented on a 15-in. (38.1-cm) screen located at eye level, 30 cm in front of participants. The same presentation sequence was used in all rounds of data collection. A brief 50-Hz orienting tone occurred, and an associated fixation cross appeared in the center of the screen 1,500 ms prior to the presentation of each stimulus. Stimuli remained on the screen until children responded. Two seconds after a response, a new orienting tone occurred, and a fixation cross appeared signaling the impending appearance of a new stimulus.

The RT familiarity task in the first phase comprised five colored pictures of familiar animals (e.g., dog, cat) and four familiar fruits (e.g., banana, apple). The task was used to introduce the stimulus presentation procedure and the speeded naming requirement. The interviewer described task requirements, encouraging children to respond as fast and as accurately as possible by naming the object that appeared on the screen (fruit or animal). None of the children exhibited difficulty understanding task requirements. The animals and the fruit naming tasks were presented separately. Each task was presented twice, with stimuli presented in a different order on each occasion.

In the DE practice task at 6 years, stimuli consisted of one to five dots in arrangements. Children completed 20 trials, comprising four different instances of each of the five dot arrangements. Children were given frequent reminders to respond as quickly and as accurately as possible; no other task-related instructions or feedback were given.

The DE experimental task comprised 40 trials of one to eight randomly arranged black dots on a white rectangular background ($n = 5$ for each numerosity). The dot stimuli presentation sequence was randomized with the constraint that not more than two trials of the same dot numerosity could follow each other. Before beginning the task, children were reminded of the procedure used in the practice session.

In the NC judgment task, children judged which of two single digit numbers was the larger in numerical magnitude. They pressed the left shift key (marked with a yellow dot) if the number on the left side of the screen was larger or the right shift key (marked with a red dot) if the number on the right side of the screen was larger. The task comprised 72 trials, representing judgment combinations of all numbers one to nine, excluding tied pairs. The larger number appeared equally often on the left side of the screen as on the right side. Four practice trials were presented initially to familiarize children with the judgment procedure.

In the naming numbers and naming letters tasks, the numbers 1–9 and the letters A–J (excluding the letter *I* because of its similarity to the number 1), respectively, were used. The two tasks comprised 36 trials, four each for the nine stimuli. The stimuli for both tasks, all of which were approximately 2 cm high on screen, were presented in one of four fixed random orders; the only constraint was that each stimulus should be different to the immediately preceding stimulus. Presentation of the naming numbers and naming letters tasks was counterbalanced.

In the basic processing speed task, children pressed a computer key as quickly as possible when a black dot appeared on the screen. The dot appeared on the screen between 500 ms and 1,000 ms after a fixation point. The task comprised nine trials.

To assess single digit addition abilities, children completed 12 single digit addition problems of the form "a + b." Addends comprised the numbers 2–7 presented in both orders (e.g., 2 + 7 and 7 + 2) and excluded tied pairs (e.g., 2 + 2). Overall accuracy was recorded. In the two-digit computation test, Children's ability to solve addition, subtraction, and multiplication problems involving up to two digits was assessed using a set of 36 problems, 12 for each operation. Problems appeared centered on a computer screen, in the form, a + b, a − b, and a × b. For addition, the addends ranged from 9 to 14; for subtraction, the minuends ranged from 11 to 16, and the subtrahends ranged from 4 to 9. The multiplication task involved multiplying two single-digit numbers from the 7, 8, and 9 times tables. In the multidigit computation test, in a single class session, children completed 36 three-digit subtraction, multiplication, and division problems (12 of each type). Three-digit addition problems were not included, as students were performing at ceiling for addition.

Measures

DE RTs were recorded by interviewers pressing a response key as soon as children gave a response. Interviewers also recorded children's answers (they were unaware of what had appeared on the computer screen). In the first phase of the study, an additional method of recording RT was used: A digital video camera focused on the computer screen, and the audio data were later analyzed using Cool Edit Pro (Syntrillium Software). The latter software, which is accurate to ±2 ms, was used to examine time sequences in the audio wave files (i.e., the onset of children's verbal responses). It was used in preference to voice-recognition timing methods because of the possibility that young children would verbalize their *thinking* prior to providing an answer, making recording accurate RT problematic. The correlation between the RT measures for the two recording methods was extremely high ($r = .99$), so for the remainder of the study, the interviewer key press was the only method used to record DE RTs (the interviewer recorded answers but was unaware of the display).

For the NC task, the ratio for each comparison was calculated by dividing the smaller number by the larger number. For example, comparing 3 with 6 gives the ratio 0.5, whereas comparing 6 with 1 gives the ratio 0.16 (i.e., 1 divided by 6). For the purpose of analysis, ratios were divided into eight ranges (i.e., .10–.19, .20–.29, . . . up to .80–.89). Ratio, rather than number distance, was used for analyses, as the research literature shows that magnitude judgments are influenced by both linear distance and the absolute magnitude of the values compared when distance is held constant (Brannon, 2006).

At each of the seven test ages, DE and NC RT data were subjected to separate Latent GOLD cluster analyses (Vermunt & Magidson, 2000, 2003). In the case of DE data, we analyzed RTs (for correct responses) for one, two, three, four, and five dots and the average of six to eight dots. For the NC data, all eight ratios were included in the analysis. Latent class cluster modeling has advantages over traditional clustering techniques in that it does not rely on traditional modeling assumptions (i.e., linear relationships, normal distributions, homogeneity), which are often violated in practice. The technique identifies subgroups by grouping people who share similar characteristics via probability-based classification. The relationship between latent classes and variables of interest can be assessed simultaneously with the identification of the clusters. This eliminates the need for the usual second stage discriminant analysis. Latent class models have recently been extended to include variables of mixed scale types (nominal, ordinal, continuous, and/or count variables) in the same analysis (Magidson & Vermunt, 2003).

For other measures (a) the simple RT measure was based on the average RT for nine problems, (b) Ravens scores were based on published norms, and (c) multidigit computation ability was based on problems solved correctly for each of the different tests: subtraction, multiplication, and division.

Results

We present findings by (a) describing general changes in DE and NC profiles over age, (b) identifying subgroups within DE and NC profiles, (c) characterizing changes over age in DE and NC subgroups, and (d) examining some relationships between DE and NC subgroups and other measures (i.e., basic processing speed, Ravens, and math ability).

Description of General DE and NC Profiles Over Age

Differences in DE RTs as a function of age and array numerosity. At all ages, children enumerated one to four dots correctly. They made more errors enumerating five or more dots at 6 years (15% overall) than at 11 years (5% overall); however, the relative error pattern remained constant across age. No RT differences were found between correct and error responses, possibly because children were allowed as long as needed to enumerate. Nevertheless, analyses reported herein were based on correct RTs only. As expected, DE RTs increased as a function of increases in dot array size and decreased as a function of age (see Table 1 for mean RTs as function of dot array numerosity and age; mean rather than median RTs are reported because they were almost

Table 1
Dot Enumeration Response Time Means and Standard Deviations as a Function of Array Numerosity and Test Age

	Test age (in years)													
	6		7		8.5		9		9.5		10		11	
Array[a]	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
1	1,401	342	1,251	192	1,088	186	1,109	246	1,073	193	1,020	175	1,012	152
2	1,585	389	1,337	289	1,080	169	1,125	251	1,068	179	1,036	205	1,007	161
3	1,919	530	1,611	349	1,267	222	1,245	228	1,170	191	1,186	229	1,181	213
4	2,660	1,070	2,030	687	1,453	374	1,424	303	1,534	364	1,547	400	1,418	300
5	3,363	1,052	2,491	680	1,901	415	2,406	656	2,093	502	1,991	586	2,070	467
6	4,907	1,723	4,259	1,457	2,129	533	2,828	672	2,855	618	2,801	811	2,574	562
7	5,519	1,850	4,886	1,702	3,376	963	3,664	998	3,477	785	3,281	670	3,205	725
8	5,572	1,829	4,510	1,316	3,037	679	3,841	905	3,654	680	3,477	877	3,458	677

[a] Number of dots in array.

identical). It is worth noting there was substantial RT variability at younger ages (see Table 1).

To assess changes in the speed of enumerating dot arrays across age, differences in RT were assessed for (a) successive numerosities at each age (e.g., three dots vs. four dots) and (b) the same dot array at contiguous test ages (e.g., 6 years vs. 7 years). The time taken to enumerate identical dot arrays decreased between 6 years and 7 years and between 7 years and 8.5 years (paired-sample t tests, $p < .05$ for all comparisons). However, after 8.5 years, decreases were smaller and often nonsignificant. Analysis of within-age changes showed that, with some exceptions (e.g., RT differences enumerating one and two dots after 8.5 years), increases in the time taken to enumerate successive dot array numerosities persisted across age, although the relative increase was less for smaller than larger dot arrays (paired-sample t tests for successive array sizes, $p < .05$ for all comparisons). These findings are consistent with previous research showing that small arrays are enumerated quickly and without error (Piazza et al., 2002; Watson et al., 2005).

Characterizing DE profiles algebraically. Relatively smaller increases in RTs occurred in the so-called subitizing range ($n \leq 4$) than in the counting range ($n \geq 4$). The change from subitizing to counting is referred to as the *point of discontinuity* (see Figures 1A and 1B). In Figure 1A, the subitizing and counting ranges are represented by separate linear functions. Figure 1B, shows the point of discontinuity represented by the change from a linear to an exponential function. A linear equation represents the best fitting RT function for one to four dots and an exponential function for one to five dots. In the latter representation, the subitizing range may be considered four dots. For linear and exponential equations, the higher the coefficient of x, the steeper the slope, and the higher the constant term (intercept), the slower the RT. We used this algebraic approach to characterize DE profiles.

The DE RT algebraic functions for age and dot numerosity are presented in Table 2. Several points are worth noting. The table shows an exponential function at 6 years for one to three dots, which suggests these children subitized two dots (i.e., RT changed in form after two dots). The table also shows an exponential function for one to four dots at 9 years, which suggests that 9-year-olds subitized three dots. The latter algebraic pattern did not change thereafter. These findings are the first to show an apparent change in the subitizing range at 7 and 9 years.

Differences in NC RTs as function of age and dot numerosity. Children made more errors and took longer to judge the larger of two Arabic numbers when they were relatively close in magnitude (e.g., 4 and 5) than when farther apart (e.g., 4 and 9). Children's accuracy was above 95% for ratios 0.10 to 0.49 (small ratios) at all ages but decreased to as low as 84% for ratios 0.50 to 0.89 (large ratios). All NC analyses reported herein were based on RTs for correct responses only.

The NC RTs for the different ratios and ages are presented in Table 3. It shows a decrease in RT for decreasing ratios and increasing age. To assess changes in the speed of NC judgments across age, differences in RT were assessed for the same ratio at contiguous test ages (e.g., 6 years vs. 7 years). The time taken to

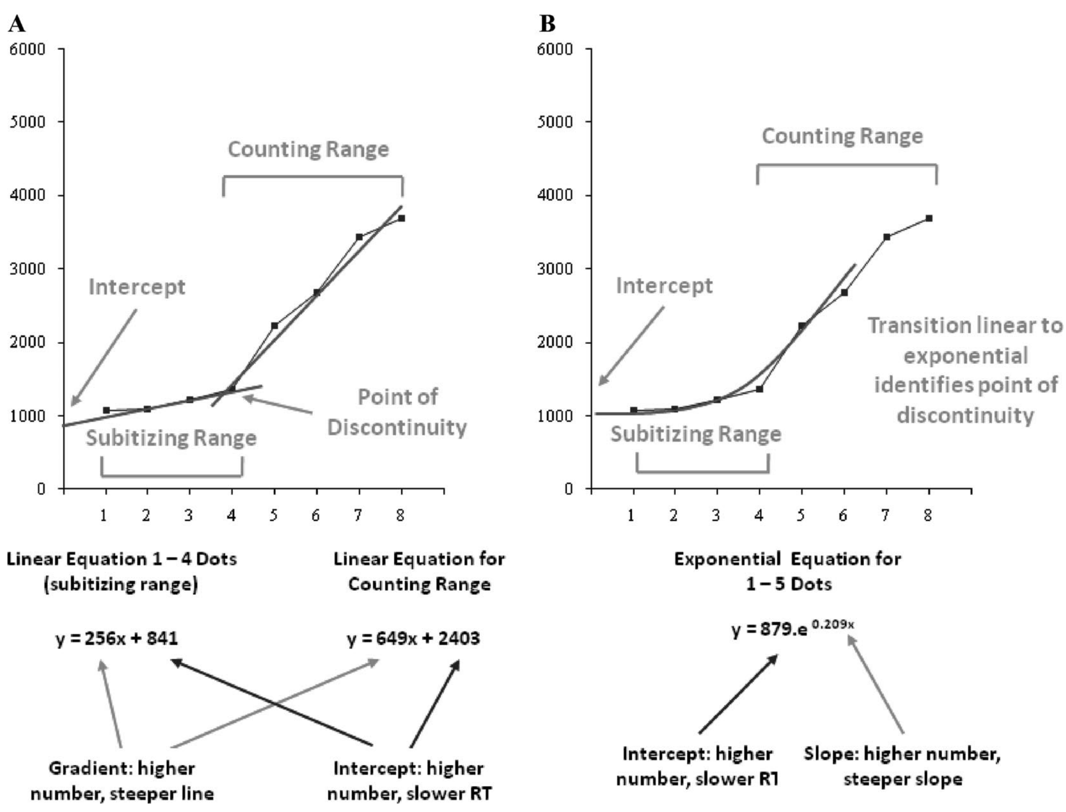

Figure 1. Example algebraic components of two 7-year-olds' dot enumeration response time (RT) profiles.

Table 2
Dot Enumeration Response Time Equations for Dot Numerosities and Age

	Dot array range			
Age (in years)	1–3 dots	1–4 dots	1–5 dots	6–8 dots
6	$y = 1{,}184e^{0.157x}$	$y = 1{,}082e^{0.212x}$	$y = 1{,}049e^{0.227x}$	$y = 724x + 3{,}031$
7	$\mathbf{y = 180x + 1{,}040}$	$y = 1{,}015e^{0.164x}$	$y = 984e^{0.180x}$	$y = 668x + 2{,}366$
8.5	$y = 90x + 966$	$y = 938e^{0.103x}$	$y = 868e^{0.141x}$	$y = 766x + 1{,}273$
9	$y = 68x + 1{,}024$	$\mathbf{y = 107x + 960}$	$y = 818e^{0.179x}$	$y = 507x + 2{,}431$
9.5	$y = 79x + 1{,}007$	$y = 148x + 840$	$y = 804e^{0.170x}$	$y = 531x + 1{,}694$
10	$y = 83x + 915$	$y = 173x + 765$	$y = 778e^{0.174x}$	$y = 494x + 1{,}653$
11	$y = 85x + 898$	$y = 139x + 807$	$y = 756e^{0.177x}$	$y = 480x + 1{,}628$

Note. Boldface indicates the point at which function changes from linear to exponential, indicating the *point of discontinuity* (i.e., a shift from subitizing to counting range).

judge each ratio decreased at successive ages between 6 years and 9 years and between 9 years and 11 years (paired-sample *t* tests, $p < .001$ for all comparisons). It should be noted that the RT standard deviations at 6 years are large but decreased with age. Changes in NC RT at each test age are expressed in terms of a function representing slope and overall speed across the ratios (see Table 4). A linear equation represents the best fitting function at all ages. It is apparent from the decreasing coefficients of *x* that the time effect of increasing ratio decreased across age. Also, successive decreases in the constant term shows that children became faster in making judgments with age.

Two points may be made about the outcomes of initial DE and NC analyses. As expected, age-related decreases in RTs were observed on both tasks; however, substantial within-age variability was also evident (see standard deviations in Tables 1 and 3). Although the algebraic equations represent these age-related changes (see Tables 2 and 4), they do not capture the variability in RTs. To determine whether the variability reflects different subgroups embedded within the overall sample, we partitioned RT data using Latent GOLD's cluster analysis program.

Identifying DE and NC Subgroups

We used Latent GOLD's cluster analysis program to identify possible subgroups embedded within the overall sample (Vermunt & Magidson, 2000, 2003; see also Notelaers, Einarsen, De Witte, and Vermunt, 2006, for use of this technique in empirical research). To determine the optimum number of clusters, three cluster models (two-, three- and four-cluster) were first identified, and the Bayesian information criteria (BIC) goodness-of-fit statistic and L^2 evaluated for each solution. If the *p* value for the four-cluster solution does not reach significance, further cluster solutions are examined. The BIC statistic is calculated by the equation 2log-likelihood + *K* log *n*, where *K* is the number of estimable parameters, and *n* is the sample size. If increasing the number of clusters results in an insignificant increase in the BIC, the smaller number of clusters is chosen. The L^2 statistic indicates the amount of association among variables that remains unexplained after estimating the model: the lower the value, the better the fit of the model to the data. One criterion for determining the number of clusters is to examine the *p* value for each model. Generally, among models in which the *p* value is greater than .05 (i.e., providing adequate fit), the one that is most parsimonious (fewest number of parameters) is selected. Alternatively, a bootstrap procedure may be used to estimate whether there is a significant difference between goodness of fit for successive cluster solutions. The latter procedure generates log-likelihood ratios and *p* values for the improvement in model fit. A nonsignificant result suggests that the subsequent model does not improve the fit of the model to the data. (We selected the Latent GOLD's cluster analysis procedure over growth curve modeling because of nonlinear changes in relationships among DE parameters over time, which would have been ignored by the latter analysis.)

Table 3
Number Comparison Response Time Means and Standard Deviations as a Function of Ratio and Test Age

	Test age (in years)													
	6		7		8.5		9		9.5		10		11	
Ratio	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
.10–.19	1,837	704	1,264	487	926	276	803	198	744	181	778	215	675	163
.20–.29	1,831	541	1,365	416	972	284	889	267	796	218	801	246	720	220
.30–.39	1,985	683	1,401	395	980	287	908	302	813	233	794	216	704	191
.40–.49	2,055	858	1,413	413	1,015	326	906	273	861	322	825	216	760	226
.50–.59	2,079	696	1,512	462	1,069	273	952	245	890	282	874	236	770	179
.60–.69	2,260	761	1,620	451	1,162	324	1,066	337	973	331	931	229	836	238
.70–.79	2,395	965	1,602	552	1,115	279	1,041	343	976	312	979	368	844	208
.80–.89	2,562	793	1,640	492	1,153	274	1,057	297	940	238	955	212	859	206

Table 4
Number Comparison Equations as a Function of Age

Age (in years)	Number comparison slope
6	y = 104x + 1,657
7	y = 54x + 1,232
8.5	y = 35x + 893
9	y = 36x + 789
9.5	y = 33x + 725
10	y = 31x + 728
11	y = 28x + 647

At each age, separate cluster analyses were conducted on the DE and NC RT data. Specifically, for DE, we analyzed RTs for one, two, three, four, and five dots (i.e., possible subitizing range) and the average of six to eight dots (counting range). Because previous studies have shown little differentiation in enumeration in the so-called counting range (i.e., >5, see Schleifer & Landerl, 2011), we thought it prudent to average performance in this range.

In the analysis of the NC data, all eight ratios were included in the analysis. On all occasions, three cluster models (two-, three- and four-cluster groups) were evaluated using the BIC (goodness-of-fit statistic and the L^2 statistics). At all seven ages, a three-cluster solution fit the DE and NC data. On the basis of changes in the BIC goodness-of-fit statistics from three- to four-cluster solutions, we are confident that a three-cluster solution best represents the age and task data ($p > .05$ in all cases). DE BIC statistics are as follows: 6 years = 1,272; 7 years = 1,210; 8.5 years = 861; 9 years = 963; 9.5 years = 934; 10 years = 864; 11 years = 855. NC BIC statistics are as follows: 6 years = 908; 7 years = 881; 8.5 years = 831; 9 years = 805; 9.5 years = 738; 10 years = 744; 11 years = 713. The variance accounted by each model in all cases was between 50% and 60%, and the best start seed was identical in replication analyses, indicating that the three-group solution was a robust representation of the data and did not represent a local maximum. Tables B1–B7 in Appendix B (see the online supplemental materials) detail the statistical measures (log-likelihood, BIC, number of parameters, and *p* values) for two-, three-, and four-cluster solutions for each phase for DE and NC.

Analyses of variance (ANOVAs) showed that the three subgroups differed in response speed profiles (described here as the fast, medium, and slow subgroups; $p < .001$ in all cases; η^2 ranged from .53 to .71). Specifically, for DE RTs at each age, the fast subgroup was faster than the medium subgroup, which, in turn, was faster than the slow subgroup ($p < .001$). The three NC subgroups were similarly systematically different from each other from 7 years onward ($p < .001$ in all cases; η^2 ranged from .55 to .75), but there were no speed differences between the subgroups at 6 years.

These findings show that similar subgroup patterns emerged from separate cluster analysis of DE and NC tasks at all ages. However, they do not show whether children remained in the same subgroup (fast, medium, or slow) across the 6 years of the study. There is no a priori reason for subgroup membership to remain stable across time.

Understanding the degree to which children's initial DE and NC cluster group assignment remained constant over time has implications for theory and practice. The next analyses examine this issue.

Consistency of DE and NC Subgroup Profiles Across Age

To characterize consistency of subgroup membership precisely, we examined membership across time in two ways. First, we computed the ordinal correlations (gamma statistics) for the DE and NC tasks. Second, we conducted a discrete factor analysis based on cluster membership assignments at each age to determine the optimum number of factors that characterize cluster solutions over time. In other words, the latter analysis assessed whether cluster membership across age aligns with factors representing distinct RT profiles.

The across-time correlations (gamma statistics) for the DE and NC subgroups are presented in Table 5.

Table 5
Ordinal Correlations (γ) for Dot Enumeration and Number Comparison Subgroups Across Age

Age (in years)	6	7	8.5	9	9.5	10	11
				Dot enumeration			
6	1.00						
7	.46**	1.00					
8.5	.48**	.59**	1.00				
9	.52**	.55**	.56**	1.00			
9.5	.57**	.55**	.64**	.59**	1.00		
10	.54**	.66**	.62**	.64**	.91**	1.00	
11	.43**	.60**	.67**	.59**	.79**	.77**	1.00
				Number comparison			
6	1.00						
7	.16	1.00					
8.5	.08	.52**	1.00				
9	.18	.53**	.53**	1.00			
9.5	.06	.43**	.69**	.61**	1.00		
10	.01	.50**	.74**	.69**	.76**	1.00	
11	.11	.53**	.47**	.50**	.54**	.54**	1.00

** $p < .01$.

It is evident from Table 5 that there is a significant ordered correlation between groupings taken at different ages. In other words, the rank ordering of subgroup membership assignment is constant across age. Unsurprisingly, the across-age DE-NC task ordinal correlations were slightly lower than the within-task correlations but were nevertheless significant (see Table 6). However, of interest is whether there is one solution for each of DE and NC that characterizes group assignment based on RT across all ages.

We used Latent GOLD discrete factor analysis to identify a solution representing group membership across time (for an example of this form of factor analysis, see Rasmussen et al., 2004). In traditional factor analysis continuous variables are expressed as a linear function of one or more continuous latent factors. Discrete factor analysis differs in several respects to traditional factor analysis: (a) Observed variables may include mixed scale types, including nominal, ordinal, continuous, and count types; (b) the latent variables are not continuous but discrete, representing two or more ordered categories; (c) the model is not linear; and (d) solutions need not be rotated to be interpreted.

A discrete factor analysis was conducted using each individual's cluster group assignment on each of the seven test occasions. A three-factor model provided the best fit for these data. For both DE and NC, the BIC statistic for each of the three-factor models was significant and the change from a three-factor to a four-factor solution did result in a better fitting model ($p > .05$ in each case; DE BIC = 2,243; NC BIC = 1,869). Moreover, the factors aligned with the RT signatures determined from cluster analyses. This result suggests that a three-factor solution provides an accurate characterization of DE and NC subgroup membership within and across age.

The subgroup solutions for DE and NC (*n*s = 30, 80, and 49 and *n*s = 39, 86, and 34, respectively) accounted for 54% and 49% of the variance in DE and NC membership. For the DE solution, the percentage of the data from each of the seven occasions contributing to the overall models ranged from 40% to 60%; and for NC, it ranged from 40% to 75%. These findings show that responses from all ages made a substantial and equal contribution to subgroup membership. Cross-classification shows a strong association between DE and NC subgroup membership, χ^2 (N = 159) = 63.05, $p < .001$; $\gamma = 0.76$, $p < .001$.

It is evident that DE and NC subgroup membership is robust over time and that overall RTs decrease over time. Of interest is whether subgroup RT changes are similar or different over time. This is issue is examined next.

Characterizing DE and NC Subgroup Changes Across Age Algebraically

Differences in DE RTs for the three cluster groups across age as a function of dot numerosity are presented in Figure 2. It is evident from Figure 2 that RTs progressively decreased across age and subgroups for each array numerosity. ANOVA of overall RT at each age revealed significant differences between the three subgroups ($p < .001$, η^2 from .42 to .60). Post hoc analyses showed that the fast subgroup enumerated all dot array numerosities faster at all ages than the medium subgroup, which, in turn, was faster than the slow subgroup. Figure 2 also illustrates that the subitizing range increases at different ages for the three subgroups.

NC RT ratio signatures for the three subgroups across age are presented in Figure 3. RT decreases across age as a function of increases in judgment ratios for each subgroup are illustrated by decreasing slopes. The average RT for the three groups systematically differed at each age from 7 years on (i.e., slow > medium > fast; $p < .001$, η^2 from .30 to .51); however, linear contrasts in repeated-measures ANOVAs showed that the gradient of the function did not differ, either within or between subgroups, across age ($p > .05$).

To characterize similarities/differences in DE and NC subgroup changes across age, we computed the same equations as reported for the overall DE and NC group equations (see Tables 2 and 4, respectively). The subgroup equations for DE and NC RTs are reported in Tables 7 and 8, respectively.

On the basis of the point at which functions change from exponential to linear (see Table 7), it is apparent that (a) the slow subgroup subitizes two dots at 6 years, three dots at 8.5 years, and four dots at 9 years; (b) the medium subgroup subitizes three dots at 6 years and four dots at 8.5 years; and (c) the fast subgroup subitizes three dots at 6 years and four dots at 7 years. The fast subgroup is consistently faster than the other two subgroups, and the medium subgroup is faster than the slow subgroup (i.e., the constant in the linear equations is greater for the slow subgroup than the other two subgroups). The slope for the subitizing range (represented by the coefficient of *x*) differs between the three groups at all test ages but does not differ in the counting range. Table 8 shows that changes in NC RTs are best described by changes in linear functions. Changes in constant terms show a decrease in RT across group and age, as do changes in the coefficient of *x*. Linear contrasts using repeated measures analysis of variance reveal no within-age differences in slope between groups at any age.

Table 6
Ordinal Correlations (γ) Between Dot Enumeration and Number Comparison Subgroups Across Age

Age (in years)	6	7	8.5	9	9.5	10	11
6	.09	.44**	.25*	.40**	.30*	.35**	.26*
7	.02	.60**	.25*	.37**	.29*	.31**	.33*
8.5	.22	.68**	.62**	.51**	.57**	.60**	.30*
9	.31*	.57**	.62**	.50**	.58**	.59**	.46**
9.5	.17	.56**	.55**	.45**	.55**	.62**	.30*
10	.29	.65**	.45**	.37**	.44**	.62**	.34*
11	.09	.54**	.47**	.36*	.41**	.43**	.48**

* $p < .05$. ** $p < .01$.

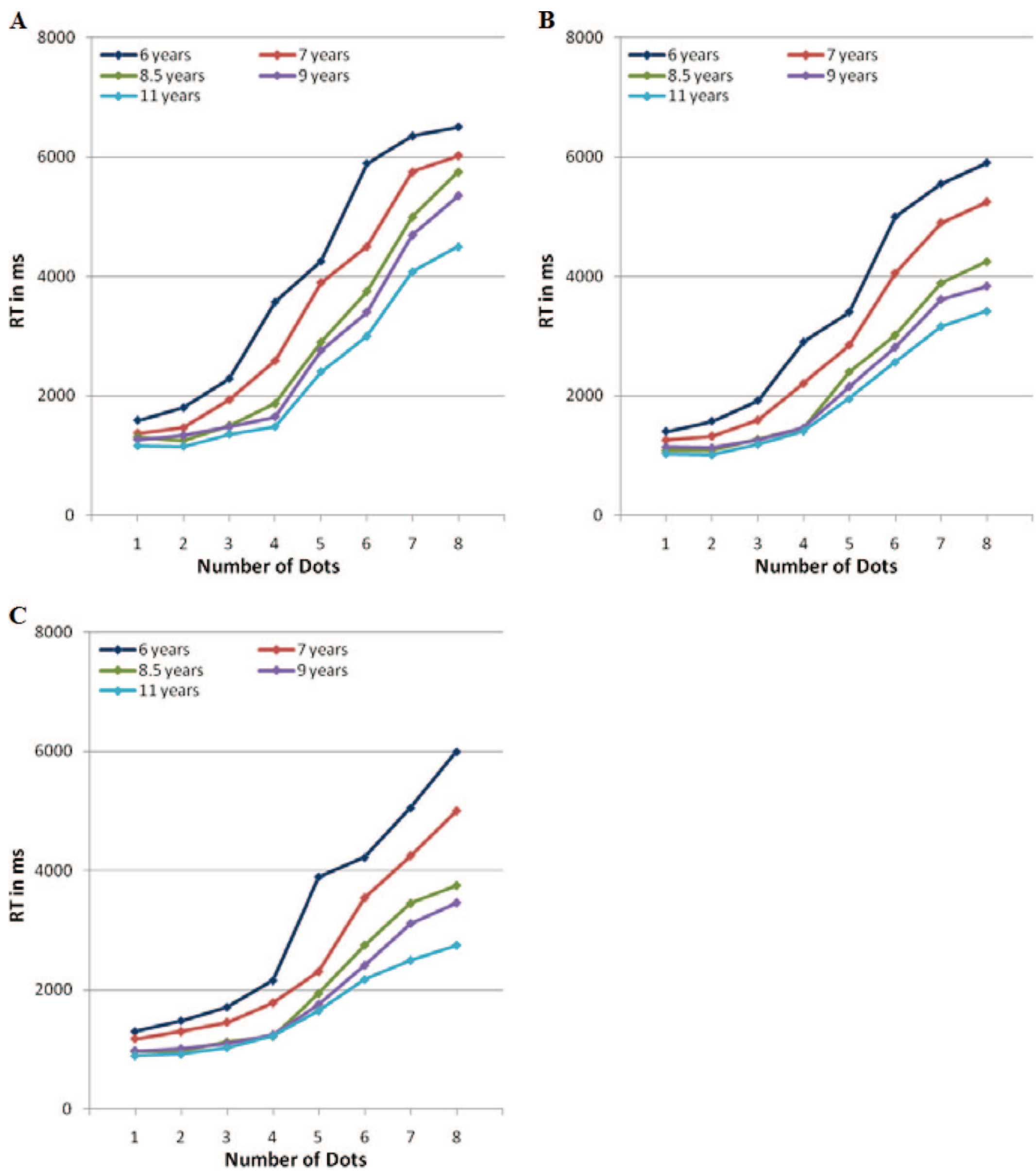

Figure 2. Mean response time (RT) dot enumeration subgroup profiles as a function of age and dot numerosity. A. Slow group. B. Medium group. C. Fast group.

Differences in DE Stable Subgroups Across Age

We focused on subgroup changes across age in DE only, because, as noted above, NC equations showed differences in RT over time but no within-age slope differences. Of interest is whether differences in the DE RT subgroup patterns represent a possible delayed or a different developmental pattern. To assess whether the slow and medium groups differed to the fast DE group, we analyzed the performance of children who were assigned to the slow, medium, or fast groups on all occasions.

Sixty-nine percent of children remained in the same subgroup across the 6 years of the study (see Figure 4).

The pattern of equations for the 109 children (see Table 11) is identical to the pattern observed for all 159 children (see Table 9). ANOVA of the average RT across all dot array numerosities (represented by the constant term in the linear functions) revealed that the slow subgroup was slower than the medium subgroup ($p < .05$); the medium subgroup was slower than the fast subgroup ($p < .05$) at all ages. Effect sizes (η^2) ranged from .40 to .57. Paired-sample *t* tests of successive ages showed that the RTs for each group decreased over time ($p < .05$).

As is evident in Table 9, the ages at which changes from exponential to linear functions occur were as follows: The slow subgroup subitized two dots at 6 years, three dots at 8.5 years, and four dots at 9 years. The medium subgroup subitized three dots initially and four dots at 8.5 years. The fast subgroup subitized three dots initially and four dots at 7 years.

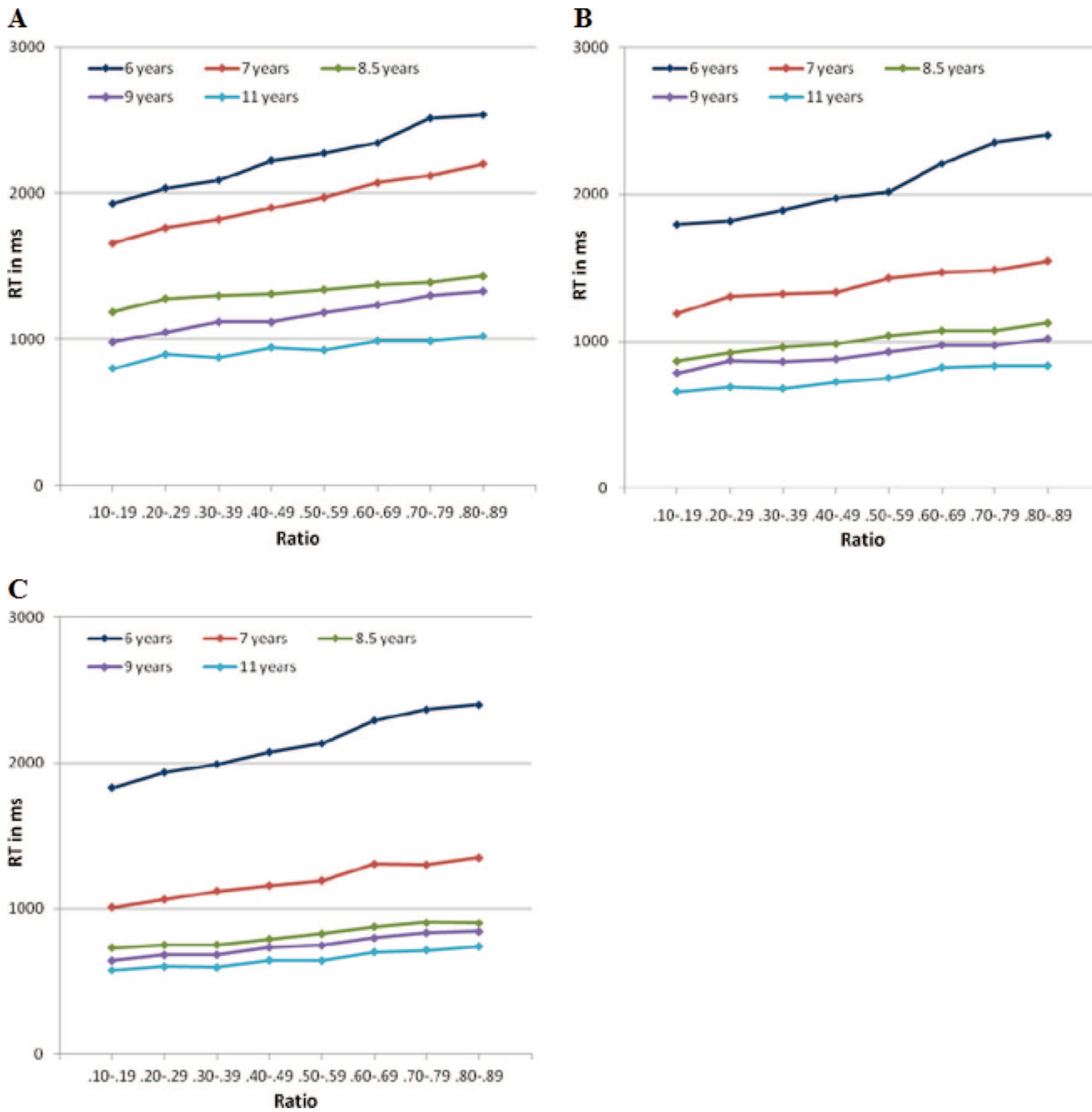

Figure 3. Mean response time (RT) number comparison subgroup profiles as a function of age and judgment ratio. A. Slow group. B. Medium group. C. Fast group.

To characterize changes in DE subgroup parameters over age, we analyzed performance at 6, 9, and 11 years (6 years = beginning of assessment, 9 years = all groups subitized to three dots, 11 years = all groups subitized to four dots). (It should be noted that no RT differences were observed in the within-age counting range gradients for the three groups.)

Repeated-measures analyses contrasting the subitizing range equations for the slow and medium subgroups indicated that the gradient of the slow subgroup was always significantly different at the three test times to the gradient of the medium subgroup: linear contrasts $F(1, 61) = 6.62$, $p < .01$, $\eta^2 = .21$ at 6 years; $F(1, 61) = 13.93$, $p < .01$, $\eta^2 = .33$ at 9 years; and $F(1, 61) = 10.54$, $p < .01$, $\eta^2 = .29$ at 11 years. However, the medium and fast subgroups did not differ in their subitizing slopes on any occasion (linear contrasts *ns*). These differences suggest that the slow DE subgroup exhibited a different subitizing pattern. Moreover, an analysis of the equations for the subitizing range of the medium and fast subgroups show that the subitizing performance of medium children at 11 years was equivalent to that of fast children at 9 years, suggesting a delay in performance.

Subgroup Differences in Simple RT, Symbol Naming RT and Nonverbal Reasoning

Because DE and NC analyses were based on RTs, it is possible that findings reflected differences in basic processing speed, symbolic access and/or general cognitive ability. To investigate these possibilities, separate one-way ANOVAs were conducted using the basic speed at 8 years and 9.5 years, RT for naming numbers and letters, and RCPM to analyze differences between the overall DE and NC subgroups (see Table 10 for values).

Table 7
Dot Enumeration Equations as a Function of Dot Numerosity, Subgroup, and Age

	Dot array range			
Group	1–3 dots	1–4 dots	1–5 dots	6–8 Dots
		6 years		
Slow	y = 1,344.e$^{0.183x}$	y = 1,165.e$^{0.268x}$	y = 1,227.e$^{0.242x}$	y = 851x + 3,508
Medium	y = 265x + 1,116	y = 1,082.e$^{0.214x}$	y = 1,037.e$^{0.235x}$	y = 746x + 3,164
Fast	y = 212x + 1,073	y = 1,058.e$^{0.175x}$	y = 996.e$^{0.205x}$	y = 647x + 2,684
		7 years		
Slow	y = 1,157.e$^{0.157x}$	y = 1,036.e$^{0.223x}$	y = 1,013.e$^{0.234x}$	y = 673x + 3,732
Medium	y = 185x + 1,035	y = 1,022.e$^{0.161x}$	y = 991.e$^{0.169x}$	y = 741x + 2,195
Fast	y = 145x + 1,028	**y = 201x + 934**	y = 972.e$^{0.154x}$	y = 577x + 2,059
		8.5 years		
Slow	**y = 101x + 1,180**	y = 1,093.e$^{0.125x}$	y = 1,037.e$^{0.151x}$	y = 667x + 1,715
Medium	y = 96x + 967	**y = 135x + 901**	y = 859.e$^{0.152x}$	y = 476x + 1,517
Fast	y = 77x + 885	y = 93x + 858	y = 819.e$^{0.119x}$	y = 376x + 1,258
		9 years		
Slow	y = 106x + 1,145	**y = 112x + 1,136**	y = 903.e$^{0.203x}$	y = 523x + 2,678
Medium	y = 62x + 1,064	y = 119x + 969	y = 833.e$^{0.183x}$	y = 535x + 1,965
Fast	y = 61x + 929	y = 89x + 882	y = 771.e$^{0.157x}$	y = 485x + 1,523
		11 years		
Slow	y = 131x + 985	y = 226x + 827	y = 816.e$^{0.218x}$	y = 534x + 2,293
Medium	y = 83x + 927	y = 134x + 843	y = 778.e$^{0.174x}$	y = 484x + 1,682
Fast	y = 69x + 829	y = 113x + 755	y = 710.e$^{0.159x}$	y = 455x + 1,308

Note. Boldface indicates the point at which the function changes from exponential to linear. An exponential function from one to four dots indicates linearity to a maximum of three dots; and an exponential function from one to three dots indicates linearity to a maximum of two dots.

Neither the analysis of RCPM—DE $F(2, 156) = 0.61$, *ns*; NC $F(2, 156) = 2.00$, *ns*—nor simple RT—DE 8 years $F(2, 156) = 2.32$, *ns*; DE 9.5 years $F(2, 156) = 1.44$, *ns*; NC 8 years $F(2, 156) = 2.47$, *ns*; NC 9.5 years $F(2, 156) = 1.97$, *ns*—revealed group effects. These findings are important because they show that the DE and NC profiles, which are based on response speeds, do not depend on differences in individuals' simple RT. In addition, one-way ANOVAs showed that neither the DE nor the NC groups differed for naming numbers or letters: DE naming numbers $F(2, 156) = 0.66$, *ns*; DE naming letters $F(2, 156) = 0.15$, *ns*; NC naming numbers $F(2, 156) = 1.23$, *ns*; NC naming letters $F(2, 156) = 2.49$, *ns*.

It should be noted that none of the RT measures were significantly associated with DE or NC group membership at any age (median F value for DE = 2.14, $p > .1$; NC = 2.35, $p > .1$).

Table 8
Number Comparison Equations as a Function of Ratio, Subgroup, and Age

Age (in years)	Slow group	Medium group	Fast group
6	y = 81x + 1,870	y = 97x + 1,609	y = 66x + 1,798
7	y = 66x + 1,610	y = 42x + 1,186	y = 41x + 977
8.5	y = 34x + 1,180	y = 33x + 843	y = 19x + 713
9	y = 44x + 968	y = 34x + 775	y = 29x + 622
11	y = 28x + 792	y = 25x + 617	y = 19x + 554

Subgroups at 6 Years and Mathematical Competence at 6, 9.5, and 11 Years

To assess whether the DE and NC subgroups were associated with math abilities, we investigated subgroup performance on single-digit addition accuracy at 6 years, two-digit arithmetic tasks at 9.5 years, and three-digit calculation tasks at 10 years (see Table 11). Only main effects were observed. The three DE subgroups all differed from each other (slow < medium < fast, $p < .05$) on single-digit addition, $F(2, 156) = 27.66$, $p < .001$, $\eta^2 = .26$. Similarly, the three NC subgroups all differed from each other (slow < medium < fast, $p < .05$) on single-digit addition, two-digit subtraction, and two-digit multiplication, $F(2, 156) = 6.95$, $p < .01$, $\eta^2 = .08$.

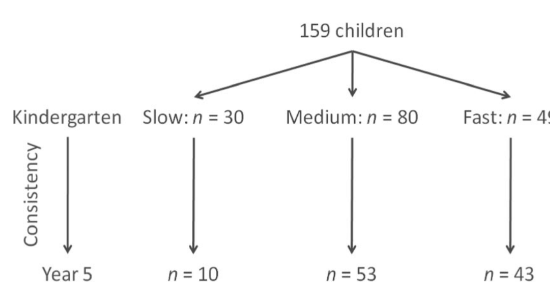

Figure 4. Consistency of dot enumeration cluster assignment from kindergarten to Year 5.

Table 9
Dot Enumeration Equations as a Function of Dot Numerosity, Stable Subgroup, and Age

Group[a]	1–3 dots	1–4 dots	1–5 dots	6–8 dots
		Dot numerosity		
		6 years		
Slow	$y = 1{,}482.e^{0.208x}$	$y = 1{,}257.e^{0.306x}$	$y = 1{,}395.e^{0.254x}$	$y = 858x + 4{,}548$
Medium	$y = 255x + 1{,}221$	$y = 1{,}143.e^{0.215x}$	$y = 1{,}085.e^{0.241x}$	$y = 753x + 3{,}375$
Fast	$y = 200x + 1{,}042$	$y = 1{,}040.e^{0.164x}$	$y = 945.e^{0.212x}$	$y = 611x + 2{,}711$
		7 years		
Slow	$y = 997.e^{0.219x}$	$y = 1{,}000.e^{0.217x}$	$y = 924.e^{0.257x}$	$y = 768x + 3{,}839$
Medium	$y = 162x + 1{,}067$	$y = 1024e^{0.158x}$	$y = 994.e^{0.173x}$	$y = 788x + 2{,}098$
Fast	$y = 136x + 1{,}040$	**$y = 193x + 946$**	$y = 971.e^{0.153x}$	$y = 567x + 2{,}029$
		8.5 years		
Slow	**$y = 154x + 1{,}080$**	$y = 1{,}018.e^{0.161x}$	$y = 1{,}024.e^{0.158x}$	$y = 652x + 1{,}726$
Medium	$y = 89x + 975$	**$y = 123x + 918$**	$y = 859e^{0.149x}$	$y = 449x + 1{,}524$
Fast	$y = 80x + 862$	$y = 90x + 845$	$y = 804.e^{0.119x}$	$y = 373x + 1{,}233$
		9 years		
Slow	$y = 124x + 1{,}335$	**$y = 160x + 1{,}309$**	$y = 982.e^{0.186x}$	$y = 565x + 2{,}714$
Medium	$y = 86x + 1{,}061$	$y = 115x + 946$	$y = 821.e^{0.178x}$	$y = 492x + 1{,}917$
Fast	$y = 78x + 909$	$y = 87x + 856$	$y = 750.e^{0.157x}$	$y = 520x + 1{,}414$
		11 years		
Slow	$y = 105x + 987$	$y = 230x + 862$	$y = 871.e^{0.202x}$	$y = 595x + 2{,}117$
Medium	$y = 83x + 9{,}041$	$y = 134x + 820$	$y = 758.e^{0.177x}$	$y = 461x + 1{,}662$
Fast	$y = 70x + 805$	$y = 113x + 734$	$y = 694.e^{0.159x}$	$y = 458x + 1{,}259$

Note. Boldface indicates the point at which the function changes from exponential to linear.
[a] Based on children who remained assigned to the same group across age.

DE subgroups all differed from each other (slow < medium < fast, $p < .05$) on two-digit addition, two-digit subtraction, and two-digit multiplication, $F(2, 156) = 19.06$, $p < .001$, $\eta^2 = .20$; $F(2, 156) = 14.63$, $p < .001$, $\eta^2 = .16$; and $F(2, 156) = 13.40$, $p < .001$, $\eta^2 = .15$, respectively. Similarly, the three NC subgroups all differed from each other (slow < medium < fast, $p < .05$) on two-digit addition, two-digit subtraction, and two-digit multiplication, $F(2, 156) = 2.95$, $p < .05$, $\eta^2 = .04$; $F(2, 156) = 3.62$, $p < .05$, $\eta^2 = .04$; and $F(2, 156) = 4.53$, $p < .05$, $\eta^2 = .06$, respectively.

Table 10
Simple RT, Symbol Naming RT and Nonverbal Reasoning Ability (RPCM) as a Function of Dot Enumeration and Number Comparison Subgroup Membership

Variable	Slow *M*	Slow *SD*	Medium *M*	Medium *SD*	Fast *M*	Fast *SD*
			Dot enumeration			
Simple RT 8 years	596.34	37.14	504.02	17.86	507.04	33.16
Simple RT 9.5 years	529.76	24.55	521.65	23.95	462.93	15.11
Name 1–9 RT	1,107.63	29.27	1,129.00	16.67	1,106.17	11.71
Name A–J[a] RT	1,458.01	55.62	1,479.34	46.92	1,452.30	28.19
RPCM	53.17	4.98	59.00	3.13	58.88	3.83
			Number comparison			
Simple RT 8 years	592.64	34.27	493.47	16.99	514.85	41.02
Simple RT 9.5 years	539.09	27.33	478.52	16.49	503.46	24.56
Name 1–9 RT	1,119.63	16.02	1,127.45	17.62	1,094.24	13.46
Name A–J[a] RT	1,457.71	31.64	1,524.27	48.31	1,402.78	33.89
RPCM	51.49	4.66	56.34	3.03	62.50	4.38

Note. RT = response time; RPCM = Ravens Colored Progressive Matrices (Raven, Court, & Raven, 1986).
[a] Excluding "I."

Table 11
Mean Percentage Correct Single-Digit Addition at 6 Years; Two-Digit Addition, Two-Digit Subtraction, and Two-Digit Multiplication at 9.5 Years; and Three-Digit Subtraction, Three-Digit Multiplication, and Three-Digit Division at 10 years as a Function of Dot Enumeration and Number Comparison Subgroup Membership

Variable	Slow		Medium		Fast	
	M	*SD*	*M*	*SD*	*M*	*SD*
Dot enumeration						
Single-digit addition	32.22	5.84	62.50	3.56	82.14	2.99
2-digit addition	71.94	5.01	90.94	1.20	92.18	1.73
2-digit subtraction	71.39	5.19	88.85	1.40	91.33	1.97
2-digit multiplication	44.44	5.08	66.67	2.72	73.64	3.39
3-digit subtraction	46.67	7.38	81.25	2.90	90.65	2.58
3-digit multiplication	60.56	6.53	85.10	2.15	87.07	3.57
3-digit division	41.67	7.02	75.62	2.88	84.86	2.97
Number comparison						
Single-digit addition	46.37	5.77	66.96	3.40	71.32	5.08
2-digit addition	82.26	3.45	88.76	1.70	91.42	2.42
2-digit subtraction	79.91	3.78	87.50	1.81	90.69	2.27
2-digit multiplication	53.85	4.53	67.34	2.66	70.10	4.66
3-digit subtraction	61.97	5.99	80.52	3.19	88.24	3.38
3-digit multiplication	73.50	5.05	82.66	2.72	85.78	3.88
3-digit division	54.06	6.02	76.16	3.07	82.35	3.37

DE subgroups all differed from each other (slow < medium < fast, $p < .05$) on three-digit subtraction, three-digit multiplication, and three-digit division, $F(2, 156) = 25.55$, $p < .001$, $\eta^2 = .25$; $F(2, 156) = 25.55$, $p < .001$, $\eta^2 = .14$; and $F(2, 156) = 25.55$, $p < .001$, $\eta^2 = .24$, respectively. Similarly, the three NC groups differed from each other (slow < medium < fast, $p < .05$) in three-digit subtraction, three-digit multiplication, and three-digit division abilities, $F(2, 156) = 9.18$, $p < .01$, $\eta^2 = .22$; $F(2, 156) = 3.55$, $p < .05$, $\eta^2 = .13$; and $F(2, 156) = 13.22$, $p < .001$, $\eta^2 = .36$, respectively. In combination, these findings suggest that differences in the DE and NC RT subgroups across age are related to math ability (but not to processing speed, symbol access, or nonverbal reasoning). It is evident that subgroup membership, which is consistent across time, predicts mathematical precocity at 9.5 and 11 years. These findings highlight the applied relevance of the identified subgroup profiles.

Discussion

This is the first longitudinal study of core numerical competencies that tracks children's arithmetic development across the entire elementary school years. Other longitudinal studies have focused on the relationship between arithmetic and cognitive abilities (e.g., Geary, 2011) or have focused on a narrower age range and more specific abilities (e.g., relationship between nonsymbolic and symbolic comparison abilities and computation between 6 and 8 years (see Desoete, Ceulemans, De Weerdt & Pieters, 2010; Stock et al., 2010). The study reported here focuses on the core indices of numerical competence of DE and NC, because these had been previously identified as cognitive markers of developmental dyscalculia (Ansari, Price, & Holloway, 2010; Bruandet et al., 2004; Butterworth, 2005a; Landerl et al., 2004; Piazza et al., 2010; Price, Holloway, Räsänen, Vesterinen, & Ansari, 2007). It was critical to establish whether these markers were stable over time and, therefore, whether they could be used in long-term prediction of children who were likely to have arithmetic learning difficulties in school. We also wished to see how these two markers were correlated, to assess the extent to which they measured the same underlying cognitive competence. Moreover, it has been argued that in DD, the principal deficit is not in understanding numerosities but, rather, linking them to their symbolic expression in words or digits (Rousselle & Noël, 2007). Thus, learners may be able to assess and compare nonsymbolic numerosities but still fail to perform normally on digit comparison (Rousselle & Noël, 2007; cf. Iuculano, Tang, Hall, & Butterworth, 2008).

Our first task was to establish in a nonarbitrary way that some children fell into distinct categories of competence. Using cluster analysis of the RTs in the DE task across the 6 years, we found that the children formed three groups (with no significant change in BIC for further groups): slow ($n = 30$), medium ($n = 80$) and fast ($n = 49$). We have been able to demonstrate that these groupings were stable from age 6 to age 11. Moreover, these groupings were significantly associated with further analyses of the RT data, including detailed analyses of slope of the RT by number of dots to be enumerated—both slope gradient and the point of discontinuity where subitizing gives way to counting. Although enumeration RTs and accuracy improved for all groups, these measures discriminate groups of children in a nonarbitrary and stable manner.

The proportion of children in the slow group at 6 years was 18%. This is higher than the modal prevalence for dyscalculia, which is around 6% (see von Aster and Shalev, 2007, for a review; see also Butterworth, 2005a; Gross-Tsur, Manor, & Shalev, 1996; Kosc, 1974). It is possible that the slow group comprises children with math learning deficit as well as those with dyscalculia (Mazzocco, 2007). However, it should be noted that the sample was not selected to be representative of the total population, because it was selected from private schools in a large city.

We used the same cluster analysis techniques in analyzing NC performance. Here we looked at both average RTs and the ratio difference between the numbers to be compared. Again we found a stable three-cluster solution characterized performance over the 6 years: slow ($n = 39$), medium ($n = 86$), and fast ($n = 34$) groups. Furthermore, the groupings for the two core competence tasks, enumeration and NC, were highly associated across all data collection rounds from 7 years onward. (It is likely that at 6 years, some children were not able to do the NC task, which would account for the high variability in RTs and accuracy in the youngest children.)

Thus, DE and NC appear to be assessing core numerical competence, which appears to be a stable individual difference. The results of this study are consistent with the hypothesis that there are stable core numerical competences (Butterworth, 1999, 2005b; Dehaene, 1997; Dehaene, Molko, & Cohen, 2004; Feigenson, Dehaene, & Spelke, 2004).

Measures of these competences were broadly stable across the 6 years of the study, even though the performance improved in all clusters, as would be expected on the basis of maturation and life experience. The finding that cluster membership remained relatively stable over time is important, given that the parameters used to identify cluster membership changed over time. In fact, 69% of children remained in the same cluster subgroup across the study, and no children changed from the medium or fast groups into the slow group. Moreover, on the basis of differences in DE parameter profiles, we suggest that the slow group differs from the medium and fast groups in processing capabilities. In particular, the RT slopes and subitizing ranges of the slow group (see Table 10) differ from those of the other two groups.

This suggestion contrasts with proposals that mathematics learning difficulties are the consequence of more general cognitive difficulties. It has been argued that a large number of "generalist genes" contribute to cognitive performance, such that a proportion of individuals will necessarily fall in the lower tail of a normal distribution on a range of educationally relevant tasks, including reading and mathematics (Kovas, Harlaar, Petrill, & Plomin, 2005). Other research has proposed that components of general cognitive ability, such as working memory, are key drivers of differences in mathematical attainment (e.g., Geary, 1993; Geary, Hoard, Nugent, & Byrd-Craven, 2007; Raghubar, Barnes, & Hecht, 2010). Rousselle and Noël (2007) have suggested that the core problem underlying arithmetical learning difficulties is an inability to link intact numerical concepts to their symbolic representations in number words or in digits. It should be noted in the present study that children in all subgroups were able to name single-digit numbers with equivalent speed and accuracy. In other words, differences in the DE and NC groups are not related to the speed of naming numbers per se. These various approaches imply that the core numerical competences measured by enumeration and comparison will bear little relationship to mathematical attainment. However, in the present research, we find that core numerical competences are related to arithmetical attainment, as has been proposed by Ansari et al. (2010); Bruandet et al. (2004); Butterworth (2003, 2005a, 2010); Butterworth, Varma, and Laurillard (2011); Koontz and Berch (1996); Landerl et al. (2004); and Piazza et al. (2010), among others. In summary, we show that DE and NC abilities are stable across age, *are not* related to general cognitive abilities, but *are* related to computation abilities.

References

Akin, O., & Chase, W. (1978). Quantification of three-dimensional structures. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 397–410. doi:10.1037/0096-1523.4.3.397

Ansari, D. (2010). Neurocognitive approaches to developmental disorders of numerical and mathematical cognition: The perils of neglecting development. *Learning and Individual Differences, 20,* 123–129. doi:10.1016/j.lindif.2009.06.001

Ansari, D., & Karmiloff-Smith, A. (2002). Atypical trajectories of number development: A neuroconstructivist perspective. *Trends in Cognitive Sciences, 6,* 511–516. doi:10.1016/S1364-6613(02)02040-5

Ansari, D., Price, G., & Holloway, I. (2010). Typical and atypical development of basic numerical magnitude representations: A review of behavioral and neuroimaging studies. In M. Ferrari & L. Vuletic (Eds.), *The developmental relations among mind, brain and education* (pp. 105–127). New York, NY: Springer. doi:10.1007/978-90-481-3666-7_5

Arp, S., & Fagard, J. (2005). What impairs subitizing in cerebral palsied children? *Developmental Psychobiology, 47,* 89–102. doi:10.1002/dev.20069

Arp, S., Taranne, P., & Fagard, J. (2006). Global perception of small numerosities (subitizing) in cerebral-palsied children. *Journal of Clinical and Experimental Neuropsychology, 28,* 405–419. doi:10.1080/13803390590935426

Bachot, J., Gevers, W., Fias, W., & Roeyers, H. (2005). Number sense in children with visuospatial disabilities: Orientation of the mental number line. *Psychology Science, 47,* 172–183.

Baker, S., Gersten, R., Flojo, J., Katz, R., Chard, D., & Clarke, B. (2002). *Preventing mathematics difficulties in young children: Focus on effective screening of early number sense delays* (Tech. Rep. No. 0305). Eugene, OR: Pacific Institutes for Research.

Balakrishnan, J. D., & Ashby, F. G. (1991). Is subitizing a unique numerical ability? *Perception & Psychophysics, 50,* 555–564. doi:10.3758/BF03207540

Balakrishnan, J. D., & Ashby, F. G. (1992). Subitizing: Magical numbers or mere superstition? *Psychological Research, 54,* 80–90. doi:10.1007/BF00937136

Basak, C., & Verhaeghen, P. (2003). Subitizing speed, subitizing range, counting speed, the Stroop effect, and aging: Capacity differences, speed equivalence. *Psychology and Aging, 18,* 240–249. doi:10.1037/0882-7974.18.2.240

Berch, D. (2005). Making sense of number sense: Implications for children with mathematical disabilities. *Journal of Learning Disabilities, 38,* 333–339. doi:10.1177/00222194050380040901

Brannon, E. M. (2006). The representation of numerical magnitude. *Current Opinion in Neurobiology, 16,* 222–229. doi:10.1016/j.conb.2006.03.002

Bruandet, M., Molko, N., Cohen, L., & Dehaene, S. (2004). A cognitive characterization of dyscalculia in Turner syndrome. *Neuropsychologia, 42,* 288–298. doi:10.1016/j.neuropsychologia.2003.08.007

Bull, R., Johnston, R. S., & Roy, J. A. (1999). Exploring the roles of the visual-spatial sketch pad and central executive in children's arithmetical skills: Views from cognition and developmental neuropsychology. *Developmental Neuropsychology, 15,* 421–442.

Butterworth, B. (1999). *The mathematical brain.* London, England: Macmillan.

Butterworth, B. (2003). *Dyscalculia screener.* London, England: NFER Nelson.

Butterworth, B. (2005a). Developmental dyscalculia. In J. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 455–467). Hove, England: Psychology Press.

Butterworth, B. (2005b). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry, 46,* 3–18. doi:10.1111/j.1469-7610.2004.00374.x

Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Sciences, 14,* 534–541. doi:10.1016/j.tics.2010.09.007

Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: From brain to education. *Science, 332,* 1049–1053. doi:10.1126/science.1201536

Chard, D., Clarke, B., Baker, S., Otterstedt, J., Braun, D., & Katz, R. (2005). Using measures of number sense to screen for difficulties in mathematics: Preliminary findings. *Assessment for Effective Intervention, 30,* 3–14. doi:10.1177/073724770503000202

Chochon, F., Cohen, L., van de Moortele, P. F., & Dehaene, S. (1999). Differential contributions of the left and right inferior parietal lobules to number processing. *Journal of Cognitive Neuroscience, 11,* 617–630. doi:10.1162/089892999563689

Cipolotti, L., Butterworth, B., & Denes, G. (1991). A specific deficit for numbers in a case of dense acalculia. *Brain, 114,* 2619–2637. doi:10.1093/brain/114.6.2619

Clarke, B., & Shinn, M. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review, 33,* 234–248.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics.* New York, NY: Oxford University Press.

Dehaene, S. (2003). The neural basis of Weber-Fechner's law: Neuronal recordings reveal a logarithmic scale for number. *Trends in Cognitive Sciences, 7,* 145–147. doi:10.1016/S1364-6613(03)00055-X

Dehaene, S., Molko, N., & Cohen, L. (2004). Arithmetic and the brain. *Current Opinion in Neurobiology, 14,* 218–224. doi:10.1016/j.conb.2004.03.008

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20,* 487–506. doi:10.1162/jocn.1993.5.4.39010.1080/02643290244000239

De Smedt, B., Swillen, A., Devriendt, K., Fryns, J. P., Verschaffel, L., & Ghesquiere, P. (2007). Mathematical disabilities in children with velocardio-facial syndrome. *Neuropsychologia, 45,* 885–895. doi:10.1016/j.neuropsychologia.2006.08.024

Desoete, A., Ceulemans, A., De Weerdt, F., & Pieters S. (2010). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *British Journal of Education Psychology.* Advance online publication. doi:10.1348/2044-8279.002002

Desoete, A., Ceulemans, A., Roeyers, H., & Huylebroeck, A. (2009). Subitizing or counting as possible screening variables for learning disabilities in mathematics education or learning? *Educational Research Review, 4,* 55–66. doi:10.1016/j.edurev.2008.11.003

Desoete, A., Roeyers, H., & De Clercq, A. (2004). Children with mathematics learning disabilities in Belgium. *Journal of Learning Disabilities, 37,* 50–61. doi:10.1177/00222194040370010601

Duncan, E., & McFarland, C. (1980). Isolating the effects of symbolic distance and semantic congruity in comparative judgments: An additive-factors analysis. *Memory & Cognition, 8,* 612–622. doi:10.3758/BF03213781

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8,* 307–314. doi:10.1016/j.tics.2004.05.002

Forster, K. I., & Forster, J. C. (2001). *DMDX* (Version 2.9.01) [Computer software]. Retrieved from http://www.u.arizona.edu/~jforster/dmdx.htm

Fuson, K. (1998). Pedagogical, mathematical, and real-world conceptual-support nets: A model for building children's multidigit domain knowledge. *Mathematical Cognition, 4,* 147–186. doi:10.1080/135467998387370

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44,* 43–74. doi:10.1016/0010-0277(92)90050-R

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

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

Geary, D. C., & Brown, S. C. (1991). Cognitive addition: Strategy choice and speed-of-processing differences in gifted, normal, and mathematically disabled children. *Developmental Psychology, 27,* 398–406.

Geary, D., Hamson, C., & Hoard, M. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology, 77,* 236–263. doi:10.1006/jecp.2000.2561

Geary, D., Hoard, M., & Hamson, C. (1999). Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology, 74,* 213–239. doi:10.1006/jecp.1999.2515

Geary, D., Hoard, M., Nugent, L., & Byrd-Craven, L. (2007). Strategy use, long-term memory, and working memory capacity. In D. Berch & M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 83–105). Baltimore, MD: Brookes.

Geary, D. C., & Lin, J. (1996). Numerical cognition: Age-related differences in the speed of executing biologically primary and biologically secondary processes. *Experimental Aging Research, 24,* 101–137.

Gersten, R., Jordan, N., & Flojo, J. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities, 38,* 293–304. doi:10.1177/00222194050380040301

Girelli, L., Lucangeli, D., & Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *Journal of Experimental Child Psychology, 76,* 104–122. doi:10.1006/jecp.2000.2564

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423,* 534–537.

Gross-Tsur, V., Manor, O., & Shalev, R. S. (1996). Developmental dyscalculia: Prevalence and demographic features. *Developmental Medicine & Child Neuropsychology, 38,* 25–33. doi:10.1111/j.1469-8749.1996.tb15029.x

Halberda, J., Mazzocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455,* 665–668. doi:10.1038/nature07246

Holloway, I. D., & 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. doi:10.1016/j.jecp.2008.04.001

Iuculano, T., Tang, J., Hall, C., & Butterworth, B. (2008). Core information processing deficits in developmental dyscalculia and low numeracy. *Developmental Science, 11,* 669–680. doi:10.1111/j.1467-7687.2008.00716.x

Jordan, N. C., Huttenlocher, J., & Levine, S. G. (1992). Differential calculation abilities in young children from middle and low-income families. *Developmental Psychology, 28,* 644–653. doi:10.1037/0012-1649.28.4.644

Jordan, N. C., Kaplan, D., Nabors Olah, L., & Locuniak, M. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development, 77,* 153–175. doi:10.1111/j.1467-8624.2006.00862.x

Kail, R. V., & Ferrer, E. (2007). Processing speed in childhood and adolescence: Longitudinal models for examining developmental change. *Child Development, 78,* 1760–1770.

Koontz, K. L., & Berch, D. B. (1996). Identifying simple numerical stimuli: Processing inefficiencies exhibited by arithmetic learning disabled children. *Mathematical Cognition, 2,* 1–24. doi:10.1080/135467996387525

Kosc, L. (1974). Developmental dyscalculia. *Journal of Learning Disabilities, 7,* 165–177. doi:10.1177/002221947400700309

Kovas, Y., Harlaar, N., Petrill, S., & Plomin, R. (2005). "Generalist genes" and mathematics in 7-year-old twins. *Intelligence, 33,* 473–489. doi:10.1016/j.intell.2005.05.002

Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: A study of 8–9- year olds. *Cognition, 93,* 99–125. doi:10.1016/j.cognition.2003.11.004

Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development, 78,* 1723–1743. doi:10.1111/j.1467-8624.2007.01087.x

Lemer, C., Dehaene, S., Spelke, E., & Cohen, L. (2003). Approximate quantities and exact number words: Dissociable systems. *Neuropsychologia, 41,* 1942–1958. doi:10.1016/S0028-3932(03)00123-4

Locuniak, M., & Jordan, N. (2008). Using kindergarten number sense to predict calculation fluency in second grade. *Journal of Learning Disabilities, 41,* 451–459. doi:0.1177/0022219408321126

Logan, G., & Zbrodoff, J. (2003). Subitizing and similarity: Toward a pattern-matching theory of enumeration. *Psychonomic Bulletin & Review, 10,* 676–682. doi:10.3758/BF03196531

Magidson, J., & Vermunt, J. K. (2003). Comparing latent class factor analysis with traditional factor analysis for data mining. In Bozdogan, H. (Ed.), *Statistical data mining and knowledge discovery* (pp. 373–383). Boca Raton, FL: CRC Press. doi:10.1201/9780203497159.ch22

Mazzocco, M. M. M. (2007). Defining and differentiating mathematical learning disabilities and difficulties. 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. 29–47). Baltimore, MD: Brookes.

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development, 82,* 1224–1237. doi:10.1111/j.1467-8624.2011.01608.x

Mazzocco, M. M. M., & Hanich, L. B. (2010). Math achievement, numerical processing, and executive functions in girls with Turner syndrome: Do girls with Turner syndrome have math learning disability? *Learning and Individual Differences, 20,* 70–81. doi:10.1016/j.lindif.2009.10.011

Mazzocco, M. M. M., & Thompson, R. (2005). Kindergarten predictors of math learning disability. *Learning Disabilities Research & Practice, 20,* 142–155. doi:10.1111/j.1540-5826.2005.00129.x

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature, 215,* 1519–1520. doi:10.1038/2151519a0

Mussolin, C., Mejias, S., & Noël, M. P. (2010). Symbolic and nonsymbolic number comparison in children with and without dyscalculia. *Cognition, 115,* 10–25. doi:10.1016/j.cognition.2009.10.006

Notelaers, G., Einarsen, S., De Witte, H., & Vermunt, J. (2006). Measuring exposure to bullying at work: The validity and advantages of the latent class cluster approach. *Work & Stress, 20,* 289–302. doi:10.1080/02678370601071594

Paterson, S., Girelli, L., Butterworth, B., & Karmiloff-Smith, A. (2006). Are numerical impairments syndrome specific? Evidence from Williams syndrome and Down's syndrome. *Journal of Child Psychology and Psychiatry, 47,* 190–204. doi:10.1111/j.1469-7610.2005.01460.x

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., . . . Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition, 116,* 33–41. doi:10.1016/j.cognition.2010.03.012

Piazza, M., Giacomini, E., Le Bihan, D., & Dehaene, S. (2003). Single-trial classification of parallel pre-attentive and serial attentive processes using functional magnetic resonance imaging. *Proceedings of the Royal Society B, 270,* 1237–1245. doi:10.1098/rspb.2003.2356

Piazza, M., Mechelli, A., Butterworth, B., & Price, C. J. (2002). Are subitizing and counting implemented as separate or functionally overlapping processes? *NeuroImage, 15,* 435–446. doi:10.1006/nimg.2001.0980

Price, G., Holloway, I., Räsänen, P., Vesterinen, M., & Ansari, D. (2007). Impaired parietal magnitude processing in developmental dyscalculia. *Current Biology, 17,* R1042–R1043. doi:10.1016/j.cub.2007.10.013

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

Rasmussen, E. R., Neuman, R. J., Heath, A. C., Levy, F., Hay, D. A., & Todd, R. D. (2004). Familial clustering of latent class and *DSM–IV* defined attention-deficit/hyperactivity disorder (ADHD) subtypes. *Journal of Child Psychology and Psychiatry, 45,* 589–598. doi:10.1111/j.1469-7610.2004.00248.x

Raven, J. C., Court, J. H., & Raven, J. (1986). *Manual for Raven's Progressive Matrices and Vocabulary Scales.* London, England: Lewis.

Rosengren, K., & Braswell, G. (2001). Variability in children's reasoning. In H. Reese & R. Kail (Eds.), *Advances in child development and behavior* (pp. 1–40). New York, NY: Academic Press.

Rousselle, L., & Noël, M. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs. nonsymbolic number magnitude. *Cognition, 102,* 361–395. doi:10.1016/j.cognition.2006.01.005

Rubinsten, O., Henik, A., Berger, A., & Shahar-Shalev, S. (2002). The development of internal representations of magnitude and their association with Arabic numerals. *Journal of Experimental Child Psychology, 81,* 74–92. doi:10.1006/jecp.2001.2645

Rueckert, L., Lange, N., Partiot, A., Appollonio, I., Litvar, I., Le Bihan, D., & Grafman, J. (1996). Visualizing cortical activation during mental calculation with functional MRI. *NeuroImage, 3,* 97–103. doi:10.1006/nimg.1996.0011

Salthouse, T. A., & Davis, H. P. (2006). Organization of cognitive abilities and neuropsychological variables across the lifespan. *Developmental Review, 26,* 31–54.

Schleifer, P., & Landerl, K. (2011). Subitizing and counting in typical and atypical development. *Developmental Science, 14,* 280–291. doi:10.1111/j.1467.2010.00976.x

Sekuler, R., & Mierkiewicz, D. (1977). Children's judgments of numerical inequality. *Child Development, 48,* 630–633. doi:10.2307/1128664

Shalev, R., Manor, O., Auerbach, J., & Gross-Tsur, V. (1998). Persistence of developmental dyscalculia: What counts? Results from a three year prospective follow-up study. *Journal of Pediatrics, 133,* 358–362. doi:10.1016/S0022-3476(98)70269-0

Shalev, R., Manor, O., & Gross-Tsur, V. (1997). Neuropsychological aspects of developmental dyscalculia. *Mathematical Cognition, 3,* 105–120. doi:10.1080/135467997387434

Shalev, R., Manor, O., & Gross-Tsur, V. (2005). Developmental dyscalculia: A prospective six-year follow-up. *Developmental Medicine & Child Neurology, 47,* 121–125. doi:10.1017/S0012162205000216

Silver, C., Pennett, D., Black, J., Fair, G., & Balise, R. (1999). Stability of arithmetic disability subtypes. *Journal of Learning Disabilities, 32,* 108–119. doi:10.1177/002221949903200202

Simon, T. J., Bearden, C. E., McDonald-McGinn, D., & Zackai, E. (2005). Visuospatial and numerical cognitive deficits in children with chromosome 22Q11.2 deletion syndrome. *Cortex, 41,* 145–155. doi:10.1016/S0010-9452(08)70889-X

Simon, T. J., Takarae, Y., DeBoer, T., McDonald-McGinn, D. M., Zackai, E. H., & Ross, J. L. (2008). Overlapping numerical cognition impairments in children with chromosome 22q11.2 deletion or Turner syndromes. *Neuropsychologia, 46,* 82–94. doi:10.1016/j.neuropsychologia.2007.08.016

Sirin, S. R. (2005). Socioeconomic status and student achievement: A meta-analytic review of research. *Review of Educational Research, 75,* 417–453. doi:10.3102/00346543075003417

Stock, P., Desoete, A., & Roeyers, H. (2010). Detecting children with arithmetic disabilities from kindergarten: Evidence from a 3-year longitudinal study on the role of preparatory arithmetic abilities. *Journal of Learning Disabilities, 43,* 250–268. doi:10.1177/0022219409345011

Svenson, O., & Sjoberg, K. (1983). Speeds of subitizing and counting processes in different age groups. *The Journal of Genetic Psychology, 142,* 203–211.

Trick, L., Enns, J., & Brodeur, D. (1996). Life span changes in visual enumeration: The number discrimination task. *Developmental Psychology, 32,* 925–932. doi:10.1037/0012-1649.32.5.925

Trick, L., & Pylyshyn, Z. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 331–351. doi:10.1037/0096-1523.19.2.331

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101,* 80–102. doi:10.1037/0033-295X.101.1.80

Tuholski, S., Engle, R., & Baylis, G. (2001). Individual differences in working memory capacity and enumeration. *Memory & Cognition, 29,* 484–492. doi:10.3758/BF03196399

Vermunt, J. K., & Magidson, J. (2000). *Latent GOLD 2.0 user's guide* [Software manual]. Belmont, MA: Statistical Innovations.

Vermunt, J. K., & Magidson, J. (2003). *Addendum to the Latent GOLD user's guide: Upgrade manual for Version 3.0* [Software manual]. Belmont, MA: Statistical Innovations.

Vetter, P., Butterworth, B., & Bahrami, B. (2011). A candidate for the attentional bottleneck: Set-size specific modulation of the right TPJ during attentive enumeration. *Journal of Cognitive Neuroscience, 23,* 728–736. doi:10.1162/jocn.2010.21472

von Aster, M., & Shalev, R. (2007). Number development and developmental dyscalculia. *Developmental Medicine & Child Neurology, 49,* 868–873. doi:10.1111/j.1469-8749.2007.00868.x

Watson, D., Maylor, E., & Bruce, L. (2005). The efficiency of feature-based subsidization and counting. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 1449–1462. doi:10.1037/0096-1523.31.6.1449

Weinert, F., & Helmke, A. (1998). The neglected role of individual differences in theoretical models of cognitive development. *Learning and Instruction, 8,* 309–323. doi:10.1016/S0959-4752(97)00024-8

Wender, K., & Rothkegel, R. (2000). Subitizing and its subprocesses. *Psychological Research, 64,* 81–92. doi:10.1007/s004260000021

Received March 30, 2011
Revision received January 27, 2012
Accepted January 28, 2012 ■