Published in final edited form as:

J Educ Psychol. 2009 May 1; 101(2): 388. doi:10.1037/a0015115.

The ABCs of Math: A Genetic Analysis of Mathematics and Its Links With Reading Ability and General Cognitive Ability

Sara A. Hart,

Department of Human Development and Family Science, Ohio State University

Stephen A. Petrill,

Department of Human Development and Family Science, Ohio State University

Lee A. Thompson, and

Department of Psychology, Case Western Reserve University

Robert Plomin

Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College, London, England

Abstract

The goal of this first major report from the Western Reserve Reading Project Math component is to explore the etiology of the relationship among tester-administered measures of mathematics ability, reading ability, and general cognitive ability. Data are available on 314 pairs of monozygotic and same-sex dizygotic twins analyzed across 5 waves of assessment. Univariate analyses provide a range of estimates of genetic ($h^2 = .00-.63$) and shared ($c^2 = .15-.52$) environmental influences across math calculation, fluency, and problem solving measures. Multivariate analyses indicate genetic overlap between math problem solving with general cognitive ability and reading decoding, whereas math fluency shares significant genetic overlap with reading fluency and general cognitive ability. Further, math fluency has unique genetic influences. In general, math ability has shared environmental overlap with general cognitive ability and decoding. These results indicate that aspects of math that include problem solving have different genetic and environmental influences than math calculation. Moreover, math fluency, a timed measure of calculation, is the only measured math ability with unique genetic influences.

Keywords

mathematics; reading; twins; genetics; environment

The National Assessment of Educational Progress (2005) survey of The Nation's Report Card reported that 64% of fourth-grade students failed to demonstrate a "proficient" level of required math skills. Moreover, the National Research Council (1999) of the National Academy of Sciences reported that very few interventions have been successful in increasing mathematics scores in low-performing children. It was further suggested in this report that the reason behind these failures is a relative lack of scientific research in mathematics ability, as well as the relationship between it and other domains such as reading and general cognitive ability.

Emerging research by Geary (2004), Jordan and colleagues (e.g., Jordan, Hanich & Kaplan, 2003), and Fuchs and colleagues (e.g., Fuchs, 2005) has begun to identify a theoretical base for understanding and measuring mathematical ability. Fuchs et al. (2005) found that nonverbal problem solving, working memory, and phonological processing differentially affected math performance depending on how it was measured (i.e., computation, concepts–applications, and fact fluency). Moreover, nonverbal reasoning, concept formation, working memory, and arithmetic number combination skill predicted computational estimation skill in third graders (Seethaler & Fuchs, 2006). In another related study of third graders, arithmetic ability was predicted by phonological decoding and processing speed, and arithmetic word problems were uniquely predicted by nonverbal problem solving, concept formation, sight word efficiency, and language (Fuchs et al., 2006). Despite this growing literature, still relatively little is known about the etiology of the relationship between mathematics, reading, and general cognitive skills (see Gersten, Jordan & Flojo, 2005, for a review).

Quantitative genetic methodology offers an important opportunity to examine the genetic and environmental etiology of the relationship between these skills. In particular, twin and adoption study methodologies allow for the examination of the proportion of variance attributable to genetic influences (or heritability; h^2), shared environmental influences (i.e., nongenetic influences that make siblings more similar; c^2), and nonshared environmental influences (i.e., nongenetic effects that make siblings different, plus error; e^2). More importantly, researchers using multivariate genetic methods can also examine genetic and environmental influences upon the covariance among mathematical ability, general cognitive ability, and reading. These methods can also be used to quantify the genetic and environmental influences on the variability in mathematics that is independent from general cognitive skills and reading.

Until recently, there have been few quantitative genetic studies examining mathematics, and most of these studies have involved participants assessed across a wide age range using somewhat limited (or different) measures of math performance. Not surprisingly, these studies have shown widely varying univariate estimates of genetic and environmental influences on math ability, from $h^2 = .20$ (Thompson, Detterman, & Plomin, 1991), to $h^2 = .66$ (Oliver et al., 2004), to $h^2 = .90$ (Alarcón, Knopik, & DeFries, 2000). The estimates of shared environmental variance are equally varied, from essentially zero (Alarcón et al., 2000; Oliver et al., 2004) to very high ($c^2 = .73$; Thompson et al., 1991).

These studies have also shown mixed results concerning the magnitude of covariance among mathematics, reading, and general cognitive skills. One approach is to assess the genetic and shared environmental correlations (r_g and r_c , respectively), which estimates the proportion of genetic and/or environmental variance that is shared across outcomes. Thompson et al. (1991) found high genetic and shared environmental correlations between math ability and reading ability ($r_g = .98$, $r_c = .92$). In an adoption study, Wadsworth, DeFries, Fulker and Plomin (1995) also examined the covariation between reading and math ability and found a genetic correlation of .80 in a parent–offspring analysis, suggesting that the genes related to math are highly correlated with those related to reading. Light and DeFries (1995) employed a related statistic to a genetic correlation, called *bivariate heritability*, or the extent to which shared genetic factors influence the overall phenotypic correlation. When examining the relationship between reading and math disability, Light and DeFries found significant effects due to bivariate heritability (.26) in young adolescent twins. Additionally, Knopik, Alarcón, and DeFries (1997) found the genetic correlation between reading and math disability to be moderate ($r_g = .53$). Most recently, Markowitz, Willemsen, Trumbetta, van Beijsterveldt, and Boomsma (2005) also found that the correlation between math problems and reading problems was almost completely explained by genetic factors ($r_g = .60$). Concerning the links between math ability and general cognitive ability, Alarcón et al. (2000) found a substantial genetic correlation between a latent factor of math ability and general cognitive ability in young

adolescents ($r_g = .95$). Interestingly, the authors also found that there were significant influences of genetics outside that shared by the two factors, suggesting unique genetic effects for both math and general cognitive ability separately. Similarly, Thompson et al. (1991) found high genetic correlations among math, verbal, and spatial ability, with no effects of shared environment ($r_g = .78-.80$).

While these previous studies have been informative, they represent findings derived from samples that were recruited across a wide age range (e.g., see Markowitz et al., 2005, where the sample was 12–24 years of age). Previous work on general cognitive ability and, more recently, reading, has suggested that genetic influences may become more important and shared environmental influences may be less important as a function of age of assessment (e.g., McGue, Bouchard, Iacono, & Lykken, 1993). Given the high degree of direct instruction in school, age-dependent shared environmental effects may become more evident if one is using a sample that is more narrowly recruited for age than previously noted in the literature containing wide age ranges. The emerging reading literature suggests that measures of reading involving direct instruction such as letter knowledge are more likely to show shared environmental effects as opposed to measures of fluency, such as rapid naming, which are highly heritable (Byrne et al., 2002; Petrill, Deater-Deckard, Thompson, DeThorne, & Schatschneider, 2006a). Thus, it also seems essential to systematically examine mathematics performance across a wide range of measures.

In response to these issues, Petrill, Plomin, and Thompson initiated an international project examining math skills. This project involves two samples, Twins Early Development Study (TEDS), an ongoing study of 5,000 pairs of twins in the United Kingdom (Oliver & Plomin, 2007), and the Western Reserve Reading Project (WRRP), an ongoing study of 314 twin pairs in the United States (Petrill, Deater-Deckard, Thompson, DeThorne, & Schatschneider, 2006b). The major goal of TEDS is to examine a large sample of twins using teacher- and Web-based measures. Analyses from TEDS have suggested consistent results for the genetic etiology of math. Oliver et al. (2004) suggested that the heritability of mathematical ability measured by teacher ratings of ability in 7-year-olds was high and significant ($h^2 = .66$). Haworth, Kovas, Petrill, and Plomin (2007) examined the same sample at 9 years old, and results suggest that the heritability estimate was similar in magnitude and significance in the slightly older children ($h^2 = .68$). Estimates of shared environment were zero and nonsignificant. Kovas, Haworth, Petrill, and Plomin (2007) measured math ability using this sample when the children were 10 years old. They found that the heritability of ability, depending on the measure, was moderate and significant ($h^2 = .42-.45$), and the shared environmental estimates were nonsignificant ($c^2 = .07-.16$). In summary, for all three articles based on the teacher ratings of math ability, genetic influences were significant, and the shared environmental estimates were low and nonsignificant.

Researchers in the TEDS project were also the first to examine the trivariate relationship between general cognitive ability, reading ability, and mathematical ability. The first such analysis was conducted when the twins were 7 years old (Kovas, Harlaar, Petrill, & Plomin, 2005), and the results suggested significant genetic correlations between general cognitive ability and mathematics ($r_g = .67$), as well as between reading and mathematics ($r_g = .74$). No significant unique influences for math were found for shared environmental effects, although a significant shared environmental correlation between reading and mathematics was found ($r_c = .74$). Similar effects were obtained when the twins were 10 years old for genetic effects underlying the relationship among math, reading, and general cognitive abilities (Davis et al., 2008). Moreover, a significant shared environmental relationship between reading and mathematics ($r_c = .94$), as well as between general cognitive ability and mathematics ($r_c = .89$), was found.

Our purpose of the present article is to present the first results on mathematics from the U.S.-based WRRP. Whereas the main strength of TEDS is the large sample size, the major advantage of WRRP is the in-depth, one-on-one, in-home assessment of mathematical and reading ability examining constructs employed in the larger math literature (e.g., Fuchs et al., 2006; Seethaler & Fuchs, 2006). The specific goals of this study were to (a) examine the univariate genetic estimates of mathematics as assessed in WRRP; (b) examine the genetic and environmental etiology of the covariance among math, reading, and general cognitive ability in WRRP; and (c) examine the etiology of the variance in math that is independent from reading and general cognitive skills. Given previous studies in related fields of specific cognitive abilities such as reading (e.g., Byrne et al., 2002; Petrill et al., 2006a), we hypothesized that univariate genetic and shared environmental estimates would be significant, although the magnitudes may differ depending on the measure, and therefore the specific skill, of math used (e.g., fluency, problem solving, etc.). With respect to multivariate genetics, we hypothesized that there would be some genetic overlap of mathematical ability with both reading ability and general cognitive ability (e.g., Plomin & Kovas, 2005) but that there will also be unique genetic effects of math alone, independent of that accounted for by reading ability and general cognitive ability.

Method

Participants

The math assessment in the WRRP is part of an ongoing longitudinal twin study in Ohio, assessed across seven home visits. Recruiting was conducted through school nominations, Ohio State birth records, and media advertisements. Schools were asked to send a packet of information to parents in their school system with twins. We secured the cooperation of 293 schools throughout the state of Ohio. We also used media advertisements in the greater Cleveland metropolitan area to recruit twins. We also hired a social worker with longstanding ties to the community to assist in the recruitment of underrepresented groups via face-to-face meetings with places of worship, community centers, and other service organizations. Three home visits were initially conducted concentrating on early reading development. A fourth visit focusing on math skills at age 8.5 years was conducted at the 6-month period before or after the third reading visit, depending on the age of the child. Recent funding allowed us to continue testing these children across three additional visits, which focused on both reading and mathematics development. Originally, 379 families showed interest in the project by enrolling, but 65 families (17%) were never tested through a home visit. This has resulted in a total sample size of 314 pairs of monozygotic (MZ; $n = 128$) and same-sex dizygotic (DZ; $n = 175$) twins with known zygosity with analyzable data (58% female twin sets). The 65 families who dropped out prior to any home visit data collection were similar in demographic background variables as the families available for analysis.

We determined twins' zygosity using DNA analysis via a cheek swab. For the cases where parents did not consent to genotyping ($n = 76$), we determined zygosity using a parent questionnaire on twins' physical similarity (Goldsmith, 1991). Although somewhat positively skewed (skew = .04), parent education levels varied widely and were similar for fathers and mothers: 12% had a high school education or less, 18% had attended some college, 30% had a bachelor's degree, 24% had some postgraduate training or a degree, and 5% did not specify. Most families were two-parent households (92%) and nearly all were White (92% of mothers, 94% of fathers).

In total, the first five waves of measurement are available for the present analyses, although data collection is ongoing for Wave 3 and 5. All twins were approximately 6 years old at Wave 1 (reading specific home visit; mean age = 6.09 years, $SD = 0.69$ years, range = 4.33–8.25yrs), approximately 7 years old at Wave 2 (reading specific home visit; mean age = 7.16 years, $SD = 0.67$ years, range = 6.00–8.83 years), approximately 8 years old at Wave 3 (reading

specific home visit; mean age = 8.32 years, $SD = 0.74$ years, range = 6.92–10.00 years), approximately 8.5 years old at Wave 4 (the math specific home visit; mean age = 8.54 years, $SD = 0.47$ years, range = 7.72–9.96 years), and approximately 10 years old for Wave 5 (math and reading specific home visit; mean age = 9.89 years, $SD = 0.80$ years, range = 8.01–12.13 years).

Procedure and Measures

General cognitive ability, mathematics performance tests of calculation, fluency, problem solving, and mathematical knowledge, as well as various reading performance tests, were administered to each twin by a separate tester. Test sessions were conducted in the twins' home in separate rooms, and the total time to complete all testing was approximately 60–90 min per child. Measures of general cognitive ability are available from Waves 1, 2, and 3 (Stanford-Binet Intelligence Scale [SB]; Thorndike, Hagen, & Sattler, 1986a). Reading measures of Rapid Automatized Naming (RAN; Wagner, Torgesen, & Rashotte, 1999) and the Word Identification and Passage Comprehension subtests of the Woodcock Reading Mastery Tests—Revised (WRMT; Woodcock, 1987) are available in Waves 1, 2, 3, and 5. All of the math measures—the Calculation, Fluency, Applied Problems, and Quantitative Concepts subtests of the Woodcock-Johnson III Tests of Achievement (WJ-III; Woodcock, McGraw, & Mather, 2001) and the Math test of the Wide Range Achievement Test—3 (WRAT; Wilkinson, 1993)—are available from Waves 4 and 5. These measures were selected for two reasons. First, they represent math outcome measures used to determine math disability status in educational settings (e.g., Fuchs et al., 2005). Second, WJ-III tests were constructed using item response theory, which will allow for better analysis of growth in math skills once additional waves of assessment are completed.

General cognitive ability—We assessed general cognitive ability using a short form of the SB (Thorndike et al., 1986a), including the Vocabulary, Pattern Analysis, Memory for Sentences, Memory for Digits, and Quantitative subtests. We summed and standardized these subtests for age and sex to form the Composite Summary of Area Score (SAS), forming a measure with a mean of 100 and a standard deviation of 15. Published internal consistency reliability of the SAS scores for the short form version of the test is high ($\alpha = .95$; Thorndike, Hagen, & Sattler, 1986b). Test-retest reliability in our sample, as measured by a Pearson product-moment correlation between the various waves of measurement, was adequate ($r_{\text{Wave 1-Wave 2}} = .70$, $r_{\text{Wave 1-Wave 3}} = .66$, $r_{\text{Wave 2-Wave 3}} = .76$). Finally, each wave of measurement loads highly onto a unitary factor of general cognitive ability, with factor loadings ranging from .76 to .96.

Reading measures—We assessed RAN using the Rapid Letter Naming and Rapid Digit Naming subtests of the Comprehensive Test of Phonological Processing (Wagner et al., 1999). Each of the subtests has a high published reliability in third graders ($\alpha_{\text{Rapid Letter Naming}} = .83$, $\alpha_{\text{Rapid Digit Naming}} = .73$; Wagner et al., 1999). Both the letter and digit subtests have been shown to be highly correlated in our sample ($r = .73$; see Petrill et al., 2006b), and therefore we calculated the mean between the two to form a RAN composite score. The composite score was reverse scored (by multiplying the mean by -1), so that a high score represented a faster naming speed. Test-retest reliability in the present sample was adequate and significant ($r = .42-.69$). Moreover, good validity was suggested in our data, as represented by each measure of RAN loading onto a single factor of RAN with high and significant factor loadings (.66–.84).

Word ID is a test of decoding from the WRMT (Woodcock, 1987), which is the ability to recognize and say out loud printed words. The individual must be able to correctly pronounce a given word. Published split-half reliability for the test is .99 for children in the third grade.

The scores for this test were standardized for age and sex normed scores, with a mean of 100 and a standard deviation of 15.

We measured Passage Comprehension via a similarly named subtest in the WRMT (Woodcock, 1987). This test is a cloze format (i.e., fill-in-the-blank) test of comprehension and has a published split-half reliability of .96 for children in the third grade. The measure was standardized for age and sex normed scores, with a mean of 100 and a standard deviation of 15.

Math measure—Calculation measures a child's ability to complete mathematical computations and is a subtest from the WJ-III (Woodcock et al., 2001). The participant must perform addition, subtraction, multiplication, division, and combinations of these. Questions contain negative numbers, percentages, decimals, fractions, and whole numbers. Published median reliability of this test is .85 in children (Woodcock et al., 2001). In the present data, test–retest reliability was moderate and significant ($r = .59$), and both measurement occasions load highly onto a unitary factor (factor loadings = .72–.79). This measure was standardized for age and sex norms, resulting in a mean of 100 and a standard deviation of 15.

Fluency measures a participant's ability to answer addition, subtraction, and multiplication problems in a 3-min time limit, and is a subtest from the WJ-III (Woodcock et al., 2001). Published median reliability of this measure is .89 in children. Test–retest reliability across the two measurements of Fluency was adequate and significant ($r = .74$), and both loaded onto one factor of Fluency (factor loadings = .84–.92). Fluency was standardized for sex and age norms, and has a mean of 100 and standard deviation of 15.

Applied Problems is a subtest of the WJ-III and measures the ability to analyze and solve applied math problems (Woodcock et al., 2001). The participant must read a problem (which includes extra, unnecessary information), decide which mathematical operation to use, and complete simple calculations using the necessary information. Published median reliability of the test in children is .92. For the present analyses, test–retest reliability between the two waves of measurement was adequate and significant ($r = .65$). Moreover, both waves load highly onto a unitary factor of Applied Problems, with factor loadings of .78 and .82. This measure was standardized for sex and age norms (resulting in a mean of 100 and standard deviation of 15).

Quantitative Concepts, also from the WJ-III, taps the knowledge of mathematical concepts, symbols, and vocabulary (Woodcock et al., 2001). There are two subtests within Quantitative Concepts: Concepts and Number Series. Concepts measures knowledge of mathematical terms and formulas, as well as the ability to count and identify numbers, shapes, and number sequences. There are no calculations required for this subtest. Number Series measures the participant's ability to find a pattern in a number series and provide a missing number to continue the series. The median reliability for the entire quantitative concepts test in children is .90. Test–retest reliability in this sample was moderate and significant ($r = .55$), and both waves loaded highly onto a single factor (factor loadings = .73–.78). The two subtests are added together, and standardized for age and sex norms, leaving a mean of 100 and standard deviation of 15.

The Math test of the WRAT assesses counting, knowledge of number symbols, solving oral problems, and doing computations. The published median reliability of the WRAT is .89 (Wilkinson, 1993). In the present data, test–retest reliability for the WRAT was significant at $r = .39$, and both measurement occasions load onto one factor, with factor loadings of .69 and .74. The WRAT was standardized for age and sex norms, resulting in a mean of 100 and a standard deviation of 15.

Results

Descriptive statistics are presented in Table 1 for general cognitive ability, reading ability, and math ability from all available waves of data collection. All measures, except for RAN, are age standardized measures around a population mean of 100 and a standard deviation of 15. In general, the sample scored around the population mean and standard deviation. As the twins represented in this sample were in different school grades (although approximately the same age), all measures collected were further corrected for months of school completed, with counting beginning the 1st month of kindergarten, prior to analysis (a Pearson product–moment correlation suggested a high degree of overlap between grade and total school months completed, $r = .81$). Additionally, for subsequent analyses, all measures were taken from raw scores and residualized for age, age squared, gender, months of school, and months of school squared, using a regression procedure. All scores were expressed in z score units following these procedures.

Creation of Composite Scores

The reading literature suggests that RAN shows discriminate validity from decoding (i.e., Hoover & Gough, 1990), so reading ability in this analysis was operationalized as a decoding score composite and RAN for more specific analysis. We created the decoding score by a mean score of the Word Identification and Passage Comprehension subtests of the WRMT (see Keenan & Betjemann, 2006, concerning the use of Passage Comprehension as a decoding measure). These two items are significantly moderately correlated in the sample within the waves ($r_{\text{Wave 1}} = .81$, $r_{\text{Wave 2}} = .88$, $r_{\text{Wave 3}} = .79$, $r_{\text{Wave 4}} = .73$) and form the Total Reading—Short Scale, as defined by Woodcock (1987). As a composite score, decoding suggested adequate test–retest stability, with moderate to high and significant correlations between the waves ($r = .54$ – $.79$). Moreover, high reliability for this composite score is suggested as all waves of measurement loaded highly onto a unitary factor of decoding, with factor loadings from .75 to .87.

Pearson Product–Moment Correlations Among Measures

Pearson product–moment correlations among general cognitive ability, reading, and math measures for all waves are presented in Table 2. General cognitive ability is significantly correlated between all waves of available data ($r = .66$ – $.76$). RAN and decoding are also significantly correlated between the waves ($r_{\text{RAN}} = .42$ – $.69$; $r_{\text{decoding}} = .54$ – $.79$). Moreover, each of the math variables are significantly correlated between the waves ($r_{\text{Calculation}} = .59$; $r_{\text{Fluency}} = .74$; $r_{\text{Applied Problems}} = .65$; $r_{\text{Quantitative Concepts}} = .55$; $r_{\text{WRAT}} = .39$). Across the measures, general cognitive ability is significantly correlated with all reading and math measures ($r = .16$ – $.70$), except for Wave 1 SB SAS with Wave 3 RAN ($r = .05$) and Wave 2 SB SAS with Wave 4 Fluency ($r = .12$). All reading variables are also significantly correlated with all available math measures ($r = .14$ – $.61$), except for Wave 2 RAN with Wave 4 Applied Problems ($r = .08$) and Wave 4 Quantitative Concepts ($r = .12$), Wave 3 RAN with Wave 4 Applied Problems ($r = .09$), and Wave 5 RAN with Wave 4 Calculation ($r = .11$), Wave 4 Applied Problems ($r = .12$), and Wave 4 Quantitative Concepts ($r = .06$).

Intraclass Twin Correlations

Next we calculated intraclass twin correlations to provide a descriptive picture of genetic and environmental influences. Results from the intraclass correlations (presented in Table 3), provide initial evidence for small to moderate univariate genetic estimates. Genetic effects are implicated because MZ correlations (.51–.87) are higher in all instances than the DZ correlations (.27–.58). Furthermore, the results suggest shared environmental effects because the MZ correlations are less than twice the DZ correlations. Finally, the intraclass correlations

suggest that there are nonshared environmental influences (and error) as the MZ twin correlations are less than unity (Neale & Cardon, 1992).

Model Fitting

Univariate analyses—Table 4 displays the results from the more specific univariate model fitting of the data for each measure at each wave. We evaluated these models in Mx on all available raw data, using 95% confidence intervals to test for significance of parameter estimates (Neale, Boker, Xie, & Maes, 2006). As was implied by the intraclass correlations above, the results indicate significant genetic influences for general cognitive ability ($h^2 = .30$ – $.55$) and all reading variables ($h^2 = .45$ – $.94$); except for Wave 5 RAN ($h^2 = .42$), as well as significant shared environmental estimates for general cognitive ability at Waves 1 and 3 ($c^2 = .34$ – $.42$). More specifically to math, Calculation in Wave 4 shows a significant shared environmental estimate only ($c^2 = .50$), whereas in Wave 5 the estimate for genetic influences is significant only ($h^2 = .44$). Math Fluency suggests significant genetic influences for Waves 4 and 5 ($h^2 = .47$ – $.63$) and shared environmental influences for Wave 5 ($c^2 = .36$). Applied Problems indicates significant genetic influences in Wave 5 only ($h^2 = .54$) and significant shared environmental influences in Wave 4 only ($c^2 = .49$). Quantitative Concepts and WRAT suggest significant shared environmental influences in Wave 4 only ($c^2 = .50$ and $c^2 = .46$, respectively). All measures across all waves show significant nonshared environmental influences, plus error ($e^2 = .12$ – $.50$).

Multivariate analyses—We used structural equation modeling to analyze the available measures to examine the genetic and environmental contributions to the variance and covariance among general cognitive ability, reading, and most importantly math ability. For the present analysis, rather than using the measured variables themselves in the model, we used a latent variable model instead. By using these latent factors, we provided evidence for construct validity of the measured variables, longitudinal stability of the measures, and provision of reliable variables (i.e., no error; see Gayan & Olson, 2003, for a review).

Therefore, the present latent variable model identified a three-factor solution from the nine measured variables (see Figure 1). Specifically, the first factor is SB SAS, which comprised Waves 1, 2, and 3 SB SAS. The second factor was reading and comprised Waves 1 and 2 RAN or Waves 1 and 2 decoding score. The third factor, math, comprised Waves 4 and 5 Calculation, Fluency, Applied Problems, Quantitative Concepts or the WRAT. This resulted in a total of 10 latent factor models being analyzed, each model representing SB SAS, one of the two reading measures (RAN or decoding) as the reading factor, and one of the five math measures (Calculation, Fluency, Applied Problems, Quantitative Concepts, or WRAT) as the math factor (see Figure 1 for model representing the model of SB SAS, RAN, and Calculation).

Along with the measurement component of the model, a Cholesky decomposition was applied to the factor solution, where the variance of the latent factors was divided into estimates of additive genetic (the combination of all alleles at all gene loci affecting the phenotype), shared environmental (all environmental influences that make family members more similar), and non-shared environmental (all environmental influences that makes family members less similar; Neale & Cardon, 1992) effects. Unlike a Cholesky decomposition on measured variables, this latent factor solution does not measure error in the nonshared environmental effects, as there is no measurement error in the factor solution.

The Cholesky decomposition (biometric) estimates the magnitude of the genetic and environmental variance components portioned into independent variances within each of the factors and shared variance between the factors. Figure 1 shows that the first set of biometric factors measure the genetic (heritability; A_1), shared environmental (C_1), and nonshared environmental (E_1) influences between all three measurement factors, or SB SAS, reading

(either RAN or decoding), and math (Calculation, Fluency, Applied Problems, Quantitative Concepts, or WRAT). The second set of biometric factors, A_2 , C_2 , and E_2 , measures the genetic and environmental influences between reading and math, outside of that explained by the SB SAS factor. Finally, the third set of biometric factors (A_3 , C_3 , and E_3) measures the unique influences on math alone, outside of the variance explained by both SB SAS and reading.

The measurement model was fit to the data using all available raw data (see Figure 1). Standardized factor loadings can be seen in Table 5 and are in general high and are all significant, suggesting a good fit for the model. More specifically, the SB SAS factor consists of three waves of SB SAS measures (factor loadings for Wave 1 = .76, for Wave 2 = .96, and for Wave 3 = .96). The reading factor is made up of Waves 1, 2, 3, and 5 RAN (factor loadings of .66, .84, .72, and .67, respectively) or decoding (factor loadings of .75, .87, .82, and .84, respectively). Finally, the math factor was formed by Waves 4 and 5 Calculation (factor loadings of .72 and .79), Fluency (factor loadings of .84 and .92), Applied Problems (factor loadings of .78 and .82), Quantitative Concepts (factor loadings of .73 and .78), or the WRAT (factor loadings of .69 and .74).

Tables 6 and 7 present the standardized path estimates from the trivariate Cholesky decomposition analyses on the measurement model (represented by Figure 1). Table 6 displays the results from these analyses with general cognitive ability, RAN, and each math factor (Calculation, Fluency, Applied Problems, Quantitative Concepts, or WRAT), resulting in five separate Cholesky factor solutions. The first set of results in Table 6 presents the outcome from the trivariate analysis between SB SAS, RAN, and Calculation (represented in Figure 1). The only significant path estimate for Calculation is represented by biometric factor C_1 and suggests that there is significant shared environmental overlap with general cognitive ability and Calculation (path estimate of .54). The second solution in Table 6 is between general cognitive ability, RAN, and Fluency. Biometric factor A_2 indicates that there is significant genetic overlap between RAN, and Fluency (path estimate of .43). Moreover, there also appears to be significant independent genetic effects for Fluency (biometric factor A_3 ; estimate of .59). The third Cholesky solution involves the math factor of Applied Problems and implies significant genetic overlap between general cognitive ability, RAN, and the math factor, represented by biometric factor A_1 (estimate of .51). Also, the biometric factor of C_1 reveals significant shared environmental overlap between general cognitive ability and Applied Problems (path estimate of .57). The fourth solution, with Quantitative Concepts as the math factor, shows genetic overlap between SB SAS, RAN, and math (biometric factor A_1 , estimate of .37), as well as a shared environmental overlap between SB SAS and math (represented by factor C_1 , estimate .71), similar to the results from when Fluency and Applied Problems are in the model. Finally, the last solution including general cognitive ability, RAN, and WRAT suggests significant genetic overlap among all three factors (biometric factor A_1 , estimate of .54) but no shared environmental overlap for WRAT. In all models presented above, none of the nonshared environmental pathways were significant for any of the math factors.

The final set of analyses, displayed in Table 7, represents five trivariate Cholesky decompositions among the factors for general cognitive ability, the decoding score (mean score of the Word Identification and Passage Comprehension subtests of the WRMT), and all of the math measures individually as the math factor. Once again, nonshared environmental covariance involving any math factor was not significant for any model.

The first solution is for the factors of general cognitive ability, the decoding score, and Calculation. Only biometric factor C_1 reveals significance for Calculation and represents the shared environmental overlap with it, general cognitive ability and decoding (path estimate of .53). The second solution, with Fluency as the math factor, implies independent genetic effects for Fluency alone, outside of general cognitive ability and the decoding score (factor

A₃, estimate of .76). The third model with Applied Problems representing the math factor suggests that there is a significant genetic overlap between general cognitive ability, the decoding score, and Applied Problems (see factor A₁), as well as a significant shared environmental overlap among all three factors (see factor C₁). The second to last model included Quantitative Concepts as the math factor and shows significant genetic and shared environmental overlap among all three factors (biometric factors A₁ and C₁, path estimates of .38 and .69, respectively). Lastly, the model with WRAT as the math factor suggests significant genetic overlap with general cognitive ability, decoding, and math (path estimate of .62).

Discussion

Our general goals for the present article were twofold. First, we presented the first univariate genetic estimates of various in-depth, tester-administered mathematical ability measures from the WRRP. Second, we conducted a multivariate analysis of math, reading, and general cognitive ability to assess the genetic and environmental overlap among the various general and specific abilities.

The univariate model fitting results for the math measures are novel in that they are the first reported estimates of genetic and environmental influences from tester-administered measures in a sample narrowly recruited at an age when formalized math instruction is beginning in earnest. Previous work on separate components of math ability from TEDS has focused on Web-based testing (e.g., Kovas et al., 2005 e.g., Kovas et al., 2007). Other researchers (e.g., Alarcón et al., 2000) have used a more limited set of math measures. We hypothesized, based on the results from other studies, that there would be moderate and significant genetic influences but that shared environmental influences may also emerge. Despite the sample differences, as well as the differences in measurement techniques from previous work in mathematics, this hypothesis was born out in the present set of analyses. For example, genetic effects were statistically significant for Fluency at both waves, and for Calculation and Quantitative Concepts at Wave 5. On the other hand, influences attributable to the shared environment were significant for Calculation, Applied Problems, Quantitative Concepts, and WRAT in Wave 4, and for only Fluency in Wave 5. In general, it appears as though there are differing significant influences of genetics and environments depending on the measurement occasion as well as the specific measure. This is especially so for the measures that suggest significant shared environmental influences, a finding slightly differing from TEDS. More specifically, despite the present shared environmental estimates being within confidence intervals of the U.K. sample, estimates from TEDS, although generally not zero in magnitude, are mostly nonsignificant. Caution must be taken when comparing the magnitude of the shared environmental estimates across studies, as power analysis suggests over 3,000 home visits would have to be conducted to test the significance of even the greatest magnitude differences between the two projects. However, the statistical significance differences between the two samples lead to some interesting hypotheses. For example, the present measures, which suggest significant shared environmental influences (e.g., Calculation) are based on skills that must be directly taught, and especially for the Wave 4 batteries, are just beginning to be taught. Therefore, these differences in significance levels may be due to educational system differences between the two countries, especially for the 1st year or two of instruction in mathematics. TEDS is based in the United Kingdom, which has a national education curriculum that every school must follow, of which there is no such system in the United States. Because both members of a twin pair typically go to the same school, variability between schools would be manifested in shared environmental estimates. Therefore, those measures that test performance on specific mathematical topics that must be directly taught would be influenced by differences in instruction across schools, possibly resulting in the significant shared environmental effects seen in the U.S. sample.

There also appears to be a trend in that the younger twins suggested more shared environmental influences for math measures that were not timed and genetic influences for the timed math measure (Fluency). On the other hand, a few years later, the univariate estimates suggest a trend toward genetic influences on all measures, and less influence of the shared environment in general. This trend corresponds with previous work on general and specific cognitive abilities, which suggest that shared environmental influences decrease as the twins experience more years of education, and subsequently genetic variance becomes more important in explaining the individual differences as variance attributable to school effects decreases (McGue et al., 1993; Pederson, Plomin, Nesselroade, & McClearn, 1992).

Beyond univariate genetic effects, the major focus of this article was to determine the overlap between math, reading, and general cognitive ability. When specific measures of both reading and math were used in the factor model, meaningful patterns emerged. This is especially so for the varying math factors, in which three distinct findings were suggested. First, when Calculation, a test of computations, was in the model as the math factor, there was no genetic overlap with general cognitive ability or reading. Instead, there was significant shared environmental overlap with general cognitive ability and decoding but not with RAN when it was in the model. This suggests that between-families environmental differences related to calculation overlap with shared environmental influences for general cognitive ability and decoding. The test of Calculation involves simple mathematical computations, a skill that must be directly learned, highlighting the importance of the environment. Further, although there is no obvious reading involved in the questions (i.e., such as in a problem solving question), there must be some general environmental influences that relate it to decoding, such as a knowledge of print or decoding the numbers.

The second major result from the multivariate analyses concerns math fluency. When RAN is in the model as the reading factor, there is a general genetic overlap between the reading and math factors. However, when decoding is the reading factor, then math fluency does not reflect shared genetic overlap, but rather independent genetic influences outside of those attributable to general cognitive ability or reading. This result suggests that the inclusion of RAN in the model results in an association between math fluency, and reading. This is not the case when RAN is not in the model, where math fluency appears to be more of an independent measure of ability. Math fluency is the only math measure presently tested with a timed component, which is similar to RAN being the only timed measure of reading ability. Perhaps it is this timed nature, rather than the actual computations, that make RAN and fluency similar to each other, reflected in the genetic overlap. In other words, it is the timed nature, not the actual specific skills, that results in genetic overlap.

The third and final set of results concern the math factors that represent math measures of problem solving. Specifically, Applied Problems, Quantitative Concepts, and the WRAT all involve reading to solve the problem, as well as deciding which information given in the question to use. For all three math problem solving measures, no matter what reading factor was in the solution, there was significant genetic overlap among general cognitive ability, reading, and math. When one examines the shared environmental influences, only Applied Problems and Quantitative Concepts suggest overlap with general cognitive ability, whereas the WRAT does not. In general, though, this pattern of genetic and environmental overlap among general cognitive ability, reading, and problem solving appears to be consistent with the fact that reading is required to answer the math problems. It can be hypothesized that the specific skills required in reading ability as well as math ability involving reading are overlapping with general cognitive ability.

In general, the results from the present multivariate analyses are similar to those obtained in TEDS (e.g., Davis et al., 2008; Kovas et al., 2005; Oliver et al., 2004). More specifically, TEDS

has shown genetic overlap, with some shared environmental overlap, between mathematics and reading, and between mathematics and general cognitive ability, which was also seen for certain components of math and reading ability. In general, the present results are similar in theme, with substantial genetic overlap suggested between general cognitive ability and math, as well as between reading and math. However, more specific conclusions could be drawn from the present study in that in-depth psychometric batteries of ability were available for analysis. The inclusion of such measures of computation, timed computation, and problem solving, with measures of reading decoding and RAN, has allowed for a more fine-tuned comparisons to be drawn, as the results described above would suggest as to the complexity of the relationship.

There are some limitations to the present study. First, due to constraints in testing time, not all measures could be given at all time points. This is a potential concern for the validity of our results. However, the moderate correlations among the measures demonstrate expected effect sizes (see Plomin & Kovas, 2005, for review). In other words, the present effect sizes are supported by the suggestion that both the magnitude of the phenotypic correlations, and the magnitude and significance of the genetic estimates, are in line with the previous literature. Another possible limitation is the high nonshared environmental univariate estimates for many of our math measures, compared to those seen in the reading ability literature. This is common across many studies of math ability, and it could be reflective of measurement error, a problem that is addressed by utilization of a common factor model, which does not include any error in the final multivariate analysis. Therefore, if the high e^2 estimates in the univariate analysis are in fact due to measurement error, then exclusion of this variance in the factor model would serve to reduce the likelihood to attaining significant results. A third limitation is that the present sample is small for some waves of analyses. Again, by utilizing a latent common factor model that takes into account all available data from all waves of analyses, we found that the power available is strengthened by capitalizing on data across waves. Moreover, in a power calculation for the present effect sizes of the univariate genetic estimates, it was determined that given our general result of MZ intraclass correlations of around .70, and the DZ intraclass correlation of around .50, we have power of .75 to reject the null hypothesis that the h^2 estimates are not different than zero. Moreover, using the same parameters, we have power of .65 to reject the null hypothesis for the c^2 estimates, and power of .99 to simultaneously reject the null for both the h^2 and c^2 estimates (Neale & Cardon, 1992). These power calculations, along with the consistency of our results with previous literature (e.g., Plomin & Kovas, 2005) and the use of a common factor model, give confidence that the present sample size is sufficient to determine significant effects. Finally, related to the small sample size limitation are the wide confidence intervals around the point estimates that are suggested by the present models. For those estimates that are significant but encompassed in a wide confidence interval, caution should be taken when interpreting the magnitude of the estimate. This is true for all quantitative genetic point estimates. Results should be held in light of significance levels, and not in the actual magnitude of the estimates, as they fall in a range of potential estimates as suggested by the confidence intervals.

The general consistencies between TEDS, powered by a large sample size using online measures of math, and WRRP, with in-depth tester administered questionnaires, is indicative that the results presented here are evocative and can be used as a welldefined base for future research. Both projects have suggested both genetic and environmental influences on math, depending on measurement. For example, the TEDS teacher-assessed and online measures of math have tended to suggest more genetic influences, whereas WRRP, while typically in the confidence intervals of TEDS, have suggested more mixed results with genetic and shared environmental influences. Despite the obvious social importance of understanding the etiology of mathematics ability in young children, previous research has fallen short in identifying the various mathematical skills that are required for success in school. The main goal of WRRP

and TEDS is to fill in this gap. The present report suggests that there is significant genetic and environmental overlap between math, reading, and general cognitive ability and that the magnitude of the overlap is dependent on which measure of mathematics is used. Over this, mathematical ability can be also be its own construct (i.e., math fluency), separate from reading and general cognitive ability by independent genes. With the longitudinal continuation of TEDS and WRRP, the growth and stability of mathematic ability and the links with reading will be more fully explored.

Acknowledgments

This work was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) Grant HD38075 well as NICHD/Department of Education Grant HD46167. We thank the twins and their families for making this research possible.

References

- Alarcón M, Knopik VS, DeFries JC. Covariation of mathematics achievement and general cognitive ability. *Journal of School Psychology* 2000;38:63–77.
- Byrne B, Delaland C, Fielding-Barnsley R, Quain P, Samuelsson S, Høien T, et al. Longitudinal twin study of early reading development in three countries: Preliminary Results. *Annals of Dyslexia* 2002;52:49–73.
- Davis OSP, Kovas Y, Harlaar N, Busfield P, McMillan A, Frances J, et al. Generalist genes and the Internet generation: Etiology of learning abilities by web testing at age 10. *Genes, Brain and Behavior* 2008;7:455–462.
- Fuchs LS. Prevention research in mathematics: Improving outcomes, building identification models, and understanding disability. *Journal of Learning Disabilities* 2005;38:350–352. [PubMed: 16122068]
- Fuchs LS, Compton DL, Fuchs D, Paulsen K, Bryant JD, Hamlett CL. The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology* 2005;97:493–513.
- Fuchs LS, Fuchs D, Compton DL, Powell SR, Seethaler PM, Capizzi AM, et al. The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology* 2006;98:29–43.
- Gayan J, Olson RK. Genetic and environmental influences on individual differences in printed word recognition. *Journal of Experimental Child Psychology* 2003;84:97–123. [PubMed: 12609495]
- Geary DC. Mathematics and learning disability. *Journal of Learning Disabilities* 2004;37:4–15. [PubMed: 15493463]
- Gersten R, Jordan NC, Flojo JR. Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities* 2005;38:293–304. [PubMed: 16122059]
- Goldsmith HH. A zygosity questionnaire for young twins: A research note. *Behavior Genetics* 1991;21:257–269. [PubMed: 1863259]
- Haworth CMA, Kovas Y, Petrill SA, Plomin R. Developmental origins of low mathematics performance and normal variation in twins from 7 to 9 years. *Twin Research and Human Genetics* 2007;10:106–117. [PubMed: 17539370]
- Hoover WA, Gough PB. The simple view of reading. *Reading and Writing* 1990;2:127–160.
- Jordan NC, Hanich LB, Kaplan D. A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with comorbid mathematics and reading difficulties. *Child Development* 2003;74:834–850. [PubMed: 12795393]
- Keenan JM, Betjemann RS. Comprehending the Gray Oral Reading Test without reading it: Why comprehension tests should not include passage-independent items. *Scientific Studies of Reading* 2006;10:363–380.
- Knopik VS, Alarcón M, DeFries JC. Comorbidity of mathematics and reading deficits: Evidence for a genetic etiology. *Behavior Genetics* 1997;27:447–453. [PubMed: 9336081]
- Kovas Y, Harlaar N, Petrill SA, Plomin R. “Generalist genes” and mathematics in 7-year-old twins. *Intelligence* 2005;33:473–489. [PubMed: 19319204]

- Kovas Y, Haworth CMA, Petrill SA, Plomin R. Mathematical ability of 10-year-old boys and girls: Genetic and environmental etiology of normal and low performance. *Journal of Learning Disabilities* 2007;40:554–567. [PubMed: 18064980]
- Light JG, DeFries JC. Comorbidity of reading and mathematics disability: Genetic and environmental etiologies. *Journal of Learning Disabilities* 1995;28:96–106. [PubMed: 7884303]
- Markowitz EM, Willemsen G, Trumbetta SL, van Beijsterveldt TCEM, Boomsma DI. The etiology of mathematical and reading (dis)ability covariation in a sample of Dutch twins. *Twin Research and Human Genetics* 2005;8:585–593. [PubMed: 16354500]
- McGue, M.; Bouchard, TJ.; Iacono, WG.; Lykken, DT. Behavioral genetics of cognitive ability: A life-span perspective. In: Plomin, R.; McClearn, GE., editors. *Nature, nurture, and psychology*. Washington, DC: American Psychological Association; 1993. p. 59–76.
- National Assessment of Educational Progress. Mathematics: National trends. 2005. Retrieved September 18, 2006, from http://nces.ed.gov/nationsreportcard/nrc/reading_math_2005/
- National Research Council. How people learn: Bridging research and practice. 1999. Retrieved September 18, 2006, from http://www.nap.edu/openbook.php?record_id_9457
- Neale, MC.; Boker, SM.; Xie, G.; Maes, HH. Mx: Statistical modeling. 7. Richmond, VA: Authors; 2006. [Computer software and manual]
- Neale, MC.; Cardon, LR. Methodology for genetic studies of twins and families. Dordrecht, the Netherlands: Kluwer Academic; 1992.
- Oliver B, Harlaar N, Hayiou-Thomas ME, Kovas Y, Walker SO, Petrill SA, et al. A twin study of teacher-reported mathematics performance and low performance in 7-year-olds. *Journal of Educational Psychology* 2004;96:504–517.
- Oliver B, Plomin R. Twins' Early Development Study (TEDS): A multivariate, longitudinal genetic investigation of language, cognition and behavior problems from childhood through adolescence. *Twin Research and Human Genetics* 2007;10:96–105. [PubMed: 17539369]
- Pederson NL, Plomin R, Nesselroade JR, McClearn GE. A quantitative genetic analysis of cognitive abilities during the second half of the life span. *Psychological Science* 1992;3:346–353.
- Petrill SA, Deater-Deckard K, Thompson LA, DeThorne LS, Schatschneider C. Genetic and environmental effects of serial naming and phonological awareness on early reading outcomes. *Journal of Educational Psychology* 2006a;98:112–121. [PubMed: 19444324]
- Petrill SA, Deater-Deckard K, Thompson LA, DeThorne LS, Schatschneider C. Reading skills in early readers: Genetic and shared environmental influences. *Journal of Learning Disabilities* 2006b;39:48–55. [PubMed: 16512082]
- Plomin R, Kovas Y. Generalist genes and learning disabilities. *Psychological Bulletin* 2005;131:592–617. [PubMed: 16060804]
- Seethaler PM, Fuchs LS. The cognitive correlates of computational estimation skill among third-grade students. *Learning Disabilities Research and Practice* 2006;21:233–243.
- Thompson LA, Detterman DK, Plomin R. Associations between cognitive abilities and scholastic achievement: Genetic overlap but environmental differences. *Psychological Science* 1991;2:158–165.
- Thorndike, RL.; Hagen, EP.; Sattler, JM. Guide for administering and scoring the fourth edition: Stanford-Binet Intelligence Scale. Chicago: Riverside; 1986a.
- Thorndike, RL.; Hagen, EP.; Sattler, JM. Technical manual. 4. Chicago: Riverside; 1986b. The Stanford-Binet Intelligence Scale.
- Wadsworth SJ, DeFries JC, Fulker DW, Plomin R. Cognitive ability and academic achievement in the Colorado Adoption Project: A multivariate genetic analysis of parent-offspring and sibling data. *Behavior Genetics* 1995;25:1–15. [PubMed: 7755514]
- Wagner, RK.; Torgesen, JK.; Rashotte, CA. Comprehensive Test of Phonological Processing. Austin, TX: PRO-ED; 1999.
- Wilkinson, GS. Wide Range Achievement Test—3. Wilmington, DE: Wide Range; 1993.
- Woodcock, RW. Woodcock Reading Mastery Tests—Revised. Circle Pines, MN: American Guidance Service; 1987.

Woodcock, RW.; McGraw, KS.; Mather, N. Woodcock-Johnson III Tests of Achievement. Itasca, IL: Riverside; 2001.

Figure 1.

Multivariate latent factor model with a Cholesky decomposition representing the overlapping genetic (A), shared environmental (C), and nonshared environmental (E) effects of Waves 1, 2, and 3 Stanford-Binet Intelligence Scale Composite Summary of Area Score factor (SB SAS); Waves 1, 2, 3, and 5 reading factor (e.g., Rapid Automated Naming [RAN]); and Waves 4 and 5 math factor (e.g., Calculation [Calc]).

Table 1

Means, Standard Deviations, Minimums, and Maximums for General Cognitive Ability, Reading Ability, and Mathematics Ability for All Waves of Available Data

Variable	M	SD	Minimum	Maximum	n
SB SAS					
Wave 1	99.92	13.03	61.00	139.00	615
Wave 2	101.25	12.40	66.00	142.00	530
Wave 3	103.29	13.51	71.00	154.00	413
RAN					
Wave 1	81.88	31.58	31.50	213.50	551
Wave 2	55.45	21.31	24.00	199.00	522
Wave 3	42.94	12.12	20.50	132.00	413
Wave 5	35.21	7.99	19.50	69.00	261
WRMT Passage Comprehension					
Wave 1	97.96	14.77	62.00	147.00	165
Wave 2	102.15	12.37	65.00	136.00	485
Wave 3	104.32	11.22	74.00	147.00	381
Wave 5	104.51	11.38	79.00	131.00	243
WRMT Word Identification					
Wave 1	104.24	18.02	76.00	174.00	597
Wave 2	111.24	12.47	81.00	151.00	520
Wave 3	109.62	10.89	83.00	137.00	410
Wave 5	107.95	10.19	80.00	130.00	238
WJ-III Calculation					
Wave 4	105.73	9.94	84.00	139.00	246
Wave 5	108.29	12.05	80.00	146.00	270
WJ-III Fluency					
Wave 4	103.37	11.42	77.00	146.00	246
Wave 5	102.11	16.78	60.00	159.00	271
WJ-III Applied Problems					
Wave 4	113.85	12.15	83.00	150.00	246
Wave 5	114.19	12.09	68.00	141.00	245

Variable	<i>M</i>	<i>SD</i>	Minimum	Maximum	<i>n</i>
WJ-III Quantitative Concepts					
Wave 4	106.30	11.04	69.00	151.00	244
Wave 5	108.21	11.02	80.00	139.00	168
WRAT					
Wave 4	101.87	11.76	74.00	136.00	216
Wave 5	101.31	15.19	76.00	150.00	90

Note. SB SAS = Stanford-Binet Intelligence Scale; RAN = Rapid Automatized Naming; WRMT = Woodcock-Reading Mastery Tests—Revised; WJ-III = Woodcock-Johnson III Tests of Achievement; WRAT = Wide Range Achievement Test—3.

Table 2

Phenotypic Correlations Between General Cognitive Ability, Reading Ability, and Math Ability for All Waves of Available Data

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
SB SAS																					
1. Wave 1	—																				
2. Wave 2	.70*	—																			
3. Wave 3	.66*	.76*	—																		
RAN																					
4. Wave 1	.23*	.26*	.26*	—																	
5. Wave 2	.19*	.31*	.25*	.57*	—																
6. Wave 3	.05	.16*	.19*	.42*	.58*	—															
7. Wave 5	.13*	.16*	.16*	.43*	.45*	.69*	—														
Decoding ^a																					
8. Wave 1	.41*	.44*	.52*	.52*	.22*	.23*	.26*	—													
9. Wave 2	.45*	.55*	.53*	.41*	.43*	.30*	.27*	.74*	—												
10. Wave 3	.46*	.47*	.54*	.32*	.31*	.34*	.25*	.66*	.68*	—											
11. Wave 5	.57*	.65*	.70*	.26*	.26*	.24*	.34*	.54*	.70*	.79*	—										
WJ-III Calculation																					
12. Wave 4	.37*	.44*	.46*	.26*	.17*	.16*	.11	.34*	.39*	.32*	.34*	—									
13. Wave 5	.36*	.39*	.40*	.32*	.25*	.15*	.18*	.43*	.43*	.42*	.45*	.59*	—								
WJ-III Fluency																					
14. Wave 4	.15*	.12	.21*	.28*	.19*	.32*	.33*	.24*	.24*	.25*	.27*	.47*	.48*	—							
15. Wave 5	.20*	.32*	.36*	.43*	.38*	.41*	.42*	.54*	.49*	.37*	.42*	.49*	.55*	.74*	—						
WJ-III Applied Problems																					
16. Wave 4	.51*	.57*	.60*	.23*	.08	.09	.12	.31*	.41*	.38*	.55*	.62*	.54*	.41*	.51*	—					
17. Wave 5	.52*	.54*	.54*	.28*	.23*	.21*	.18*	.51*	.56*	.56*	.62*	.51*	.67*	.49*	.50*	.65*	—				
WJ-III Quantitative Concepts																					
18. Wave 4	.49*	.50*	.54*	.20*	.12	.14*	.06	.36*	.50*	.41*	.49*	.54*	.45*	.34*	.36*	.57*	.57*	—			
19. Wave 5	.40*	.47*	.49*	.26*	.22*	.20*	.22*	.61*	.54*	.50*	.57*	.49*	.65*	.49*	.48*	.57*	.71*	.55*	—		

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
WRAT																					
20. Wave 4	.40*	.43*	.50*	.25*	.25*	.10	.09	.40*	.43*	.31*	.34*	.73*	.54*	.43*	.46*	.57*	.46*	.56*	.35*	—	—
21. Wave 5	.37*	.43*	.41*	.32*	.20*	.21	.25*	.49*	.51*	.46*	.54*	.45*	.72*	.52*	.57*	.51*	.61*	.32*	.60*	.39*	—

Note. SB SAS = Stanford-Binet Intelligence Scale; RAN = Rapid Automatized Naming; WJ-III = Woodcock-Johnson III Tests of Achievement; WRAT = Wide Range Achievement Test—3.

^aMean score of Word Identification and Passage Comprehension subtests of the Woodcock Reading Mastery Tests—Revised.

*
p = .05

Table 3

Twin Intraclass Correlations Between General Cognitive Ability, Reading Ability and Mathematics Ability For All Waves of Available Data

Variable	Twin intraclass correlations	
	MZ	DZ
SB SAS		
Wave 1	.76*	.55*
Wave 2	.77*	.46*
Wave 3	.65*	.46*
RAN		
Wave 1	.71*	.33*
Wave 2	.56*	.43*
Wave 3	.68*	.27*
Wave 5	.54*	.42*
Decoding ^a		
Wave 1	.82*	.50*
Wave 2	.86*	.44*
Wave 3	.77*	.40*
Wave 5	.85*	.29*
WJ-III Calculation		
Wave 4	.53*	.51*
Wave 5	.73*	.50*
WJ-III Fluency		
Wave 4	.82*	.41*
Wave 5	.87*	.58*
WJ-III Applied Problems		
Wave 4	.59*	.56*
Wave 5	.73*	.51*
WJ-III Quantitative Concepts		
Wave 4	.51*	.49*
Wave 5	.75*	.49*
WRAT		
Wave 4	.58*	.45*
Wave 5	.66*	.48*

Note. MZ = monozygotic; DZ = same-sex dizygotic; SB SAS = Stanford-Binet Intelligence Scale; RAN = Rapid Automatized Naming; WJ-III = Woodcock-Johnson III Tests of Achievement; WRAT = Wide Range Achievement Test—3.

^a Mean score of Word Identification and Passage Comprehension subtests of the Woodcock Reading Mastery Tests—Revised.

* $p = .05$.

Table 4

Univariate Estimates for Heritability (h^2), Shared Environment (c^2), and Nonshared Environmental (e^2) Influences for All Measures of General Cognitive Ability, Reading Ability, and Mathematics Ability

Variable	h^2	c^2	e^2
SB SAS			
Wave 1	.30 (.09–.53)	.42 (.19–.63)	.25 (.20–.33)
Wave 2	.55 (.29–.84)	.21 (.00–.47)	.25 (.19–.33)
Wave 3	.36 (.01–.74)	.34 (.01–.65)	.33 (.24–.47)
RAN			
Wave 1	.72 (.42–.88)	.00 (.00–.27)	.29 (.22–.39)
Wave 2	.79 (.52–.96)	.00 (.00–.00)	.22 (.17–.31)
Wave 3	.65 (.49–.83)	.00 (.00–.00)	.32 (.24–.44)
Wave 5	.42 (.00–.82)	.17 (.00–.55)	.39 (.28–.59)
Decoding ^a			
Wave 1	.45 (.11–.87)	.28 (.00–.65)	.19 (.13–.30)
Wave 2	.76 (.52–.98)	.07 (.00–.33)	.15 (.11–.20)
Wave 3	.81 (.56–1.00)	.00 (.00–.00)	.17 (.12–.24)
Wave 5	.94 (.72–1.20)	.00 (.00–.00)	.12 (.08–.18)
WJ-III Calculation			
Wave 4	.04 (.00–.52)	.50 (.10–.78)	.44 (.30–.60)
Wave 5	.44 (.10–.84)	.29 (.00–.63)	.26 (.18–.38)
WJ-III Fluency			
Wave 4	.63 (.27–.98)	.15 (.00–.53)	.21 (.14–.32)
Wave 5	.47 (.23–.78)	.36 (.04–.66)	.14 (.10–.20)
WJ-III Applied Problems			
Wave 4	.14 (.00–.58)	.49 (.12–.82)	.37 (.25–.54)
Wave 5	.54 (.17–.95)	.24 (.00–.60)	.22 (.15–.35)
WJ-III Quantitative Concepts			
Wave 4	.00 (.00–.00)	.50 (.07–.66)	.50 (.35–.66)
Wave 5	.29 (.00–.92)	.52 (.00–.99)	.23 (.14–.40)
WRAT			
Wave 4	.13 (.00–.64)	.46 (.02–.81)	.42 (.28–.60)
Wave 5	.35 (.00–.93)	.34 (.00–.85)	.32 (.21–.54)

Note. The 95% confidence intervals are in parentheses. SB SAS = Stanford-Binet Intelligence Scale; RAN = Rapid Automatized Naming; WJ-III = Woodcock-Johnson III Tests of Achievement; WRAT = Wide Range Achievement Test—3.

^aMean score of Word Identification and Passage Comprehension subtests of the Woodcock Reading Mastery Tests—Revised.

Table 5

Factor Loadings for All Measures for the SB SAS Factor (Waves 1, 2, and 3 Stanford-Binet Intelligence Scale), Reading Factor (RAN or Decoding), and the Math Factor (WJ-III Calculation, Fluency, Applied Problems, Quantitative Concepts, or WRAT)

Variable	SB SAS factor	RAN factor	Reading (Decoding) factor	Math (WJ-III Calculation) factor	Math (WJ-III Fluency) factor	Math (WJ-III Applied Problems) factor	Math (WJ-III Quantitative Concepts) factor	Math (WRAT) factor
SB SAS								
Wave 1	.76							
Wave 2	.96							
Wave 3	.96							
RAN								
Wave 1		.66						
Wave 2		.84						
Wave 3		.72						
Wave 5		.67						
Decoding ^a								
Wave 1			.75					
Wave 2			.87					
Wave 3			.82					
Wave 5			.84					
WJ-III Calculation								
Wave 4				.72				
Wave 5				.79				
WJ-III Fluency								
Wave 4					.84			
Wave 5					.92			
WJ-III Applied Problems								
Wave 4						.78		
Wave 5						.82		
WJ-III Quantitative Concepts								
Wave 4							.73	
Wave 5							.78	
WRAT								

Variable	SB SAS factor	RAN factor	Reading (Decoding) factor	Math (WJ-III Calculation) factor	Math (WJ-III Fluency) factor	Math (WJ-III Applied Problems) factor	Math (WJ-III Quantitative Concepts) factor	Math (WRAT) factor
Wave 4								.69
Wave 5								.74

Note. SB SAS = Stanford-Binet Intelligence Scale; RAN = Rapid Automatized Naming; WJ-III = Woodcock-Johnson III Tests of Achievement; WRAT = Wide Range Achievement Test—3.

^aMean score of Word Identification and Passage Comprehension subtests of the Woodcock Reading Mastery Tests—Revised.

Table 6

Multivariate Modeling Results of Genetic and Environmental Influences on Stanford-Binet Intelligence Scale SAS Factor, the Reading Factor (RAN), and the Math Factor (Calculation, Fluency, Applied Problems, Quantitative Concepts, or WRAT)

Variable	Overlap between SB SAS, reading, and math	Overlap between reading and math	Unique effects of math
	A ₁	A ₂	A ₃
SB SAS	.73 (.56–.89)		
RAN	.31 (.06–.56)	.90 (.75–.95)	
WJ-III Calculation	.30 (.00–.64)	.13 (.00–.33)	.16 (.00–.64)
	C ₁	C ₂	C ₃
SB SAS	.56 (.25–.72)		
RAN	.19 (.00–.46)	.05 (.00–.45)	
WJ-III Calculation	.54 (.08–.98)	.75 (.00–.88)	.00 (.00–.88)
	E ₁	E ₂	E ₃
SB SAS	.40 (.33–.48)		
RAN	.13 (.01–.26)	.19 (.00–.36)	
WJ-III Calculation	.13 (.00–.33)	.00 (.00–.33)	.00 (.00–.34)
	A ₁	A ₂	A ₃
SB SAS	.73 (.56–.89)		
RAN	.33 (.06–.57)	.89 (.73–.94)	
WJ-III Fluency	.20 (.00–.46)	.43 (.20–.61)	.59 (.37–.80)
	C ₁	C ₂	C ₃
SB SAS	.56 (.23–.72)		
RAN	.17 (.00–.46)	.05 (.00–.47)	
WJ-III Fluency	.27 (.00–.71)	.58 (.00–.77)	.00 (.00–.73)
	E ₁	E ₂	E ₃
SB SAS	.40 (.33–.48)		
RAN	.13 (.01–.25)	.22 (.00–.37)	
WJ-III Fluency	.00 (.00–.06)	.10 (.00–.28)	.00 (.00–.24)
	A ₁	A ₂	A ₃
SB SAS	.73 (.56–.89)		
RAN	.36 (.09–.58)	.89 (.73–.94)	
WJ-III Applied Problems	.51 (.17–.84)	.00 (.00–.20)	.45 (.00–.73)
	C ₁	C ₂	C ₃

Variable	Overlap between SB SAS, reading, and math	Overlap between reading and math	Unique effects of math
SB SAS	.56 (.23–.72)		
RAN	.13 (.00–.42)	.00 (.00–.46)	
WJ-III Applied Problems	.57 (.12–.89)	.00 (.00–.70)	.43 (.00–.70)
	E ₁	E ₂	E ₃
SB SAS	.40 (.33–.48)		
RAN	.13 (.01–.25)	.22 (.00–.37)	
WJ-III Applied Problems	.04 (.00–.24)	.16 (.00–.39)	.00 (.00–.31)
	A ₁	A ₂	A ₃
SB SAS	.73 (.56–.88)		
RAN	.34 (.09–.56)	.89 (.75–.95)	
WJ-III Quantitative Concepts	.37 (.01–.77)	.06 (.00–.30)	.45 (.00–.75)
	C ₁	C ₂	C ₃
SB SAS	.56 (.28–.72)		
RAN	.17 (.00–.44)	.00 (.00–.45)	
WJ-III Quantitative Concepts	.71 (.21–.98)	.02 (.00–.73)	.33 (.00–.73)
	E ₁	E ₂	E ₃
SB SAS	.39 (.33–.47)		
RAN	.13 (.01–.25)	.21 (.00–.37)	
WJ-III Quantitative Concepts	.06 (.00–.29)	.21 (.00–.48)	.00 (.00–.42)
	A ₁	A ₂	A ₃
SB SAS	.73 (.56–.89)		
RAN	.32 (.07–.57)	.90 (.74–.95)	
WRAT	.54 (.19–.87)	.16 (.00–.35)	.00 (.00–.51)
	C ₁	C ₂	C ₃
SB SAS	.56 (.25–.72)		
RAN	.19 (.00–.47)	.00 (.00–.45)	
WRAT	.31 (.00–.74)	.00 (.00–.86)	.76 (.00–.86)
	E ₁	E ₂	E ₃
SB SAS	.40 (.33–.47)		
RAN	.13 (.01–.25)	.19 (.00–.36)	
WRAT	.10 (.00–.31)	.07 (.00–.33)	.00 (.00–.29)

Note. The 95% confidence intervals are in parentheses. The first set of biometric factors measure the genetic (A₁), shared environmental (C₁), and nonshared environmental (E₁) influences between all three measurement factors. The second set (A₂, C₂, and E₂) measures the genetic and environmental influences between reading and math, outside of that explained by the SB SAS factor. Finally, the third set of biometric factors (A₃,

C₃, and E₃) measures the unique influences on math alone, outside of the variance explained by both SB SAS and reading. SB SAS = Stanford-Binet Intelligence Scale; RAN = Rapid Automatized Naming; WJ-III = Woodcock-Johnson III Tests of Achievement; WRAT = Wide Range Achievement Test—3.

Table 7

Multivariate Modeling Results of Genetic and Environmental Influences on Stanford-Binet Intelligence Scale SAS Factor, the Reading Factor (Decoding Score), and the Math Factor (Calculation, Fluency, Applied Problems, Quantitative Concepts, or WRAT)

Variable	Overlap between SB SAS, reading, and math	Overlap between reading and math	Unique effects of math
	A ₁	A ₂	A ₃
SB SAS	.73 (.56–.89)		
Decoding score ^a	.64 (.46–.82)	.68 (.51–.76)	
WJ-III Calculation	.29 (.00–.63)	.29 (.00–.54)	.00 (.00–.61)
	C ₁	C ₂	C ₃
SB SAS	.55 (.23–.71)		
Decoding score	.36 (.07–.55)	.01 (.00–.30)	
WJ-III Calculation	.53 (.06–.95)	.73 (.00–.88)	.00 (.00–.86)
	E ₁	E ₂	E ₃
SB SAS	.40 (.33–.48)		
Decoding score	.06 (.00–.16)	.00 (.00–.19)	
WJ-III Calculation	.13 (.00–.31)	.00 (.00–.35)	.00 (.00–.34)
	A ₁	A ₂	A ₃
SB SAS	.73 (.56–.90)		
Decoding score	.62 (.44–.82)	.67 (.48–.76)	
WJ-III Fluency	.15 (.00–.43)	.24 (.00–.50)	.76 (.59–.93)
	C ₁	C ₂	C ₃
SB SAS	.55 (.22–.71)		
Decoding score	.37 (.09–.57)	.16 (.00–.36)	
WJ-III Fluency	.26 (.00–.70)	.53 (.00–.75)	.00 (.00–.63)
	E ₁	E ₂	E ₃
SB SAS	.40 (.33–.48)		
Decoding score	.06 (.00–.16)	.00 (.00–.20)	
WJ-III Fluency	.00 (.00–.63)	.00 (.00–.24)	.00 (.00–.24)
	A ₁	A ₂	A ₃
SB SAS	.73 (.55–.89)		
Decoding score	.64 (.46–.83)	.68 (.50–.76)	
WJ-III Applied Problems	.56 (.25–.85)	.24 (.00–.47)	.45 (.00–.71)
	C ₁	C ₂	C ₃

Variable	Overlap between SB SAS, reading, and math	Overlap between reading and math	Unique effects of math
SB SAS	.56 (.25–.72)		
Decoding score	.35 (.07–.55)	.00 (.00–.30)	
WJ-III Applied Problems	.51 (.08–.84)	.12 (.00–.68)	.39 (.00–.67)
	E ₁	E ₂	E ₃
SB SAS	.40 (.33–.48)		
Decoding score	.06 (.00–.16)	.00 (.00–.18)	
WJ-III Applied Problems	.03 (.00–.18)	.00 (.00–.32)	.00 (.00–.31)
	A ₁	A ₂	A ₃
SB SAS	.73 (.56–.88)		
Decoding score	.64 (.46–.82)	.68 (.51–.76)	
WJ-III Quantitative Concepts	.38 (.05–.77)	.40 (.08–.67)	.30 (.00–.68)
	C ₁	C ₂	C ₃
SB SAS	.55 (.26–.71)		
Decoding score	.35 (.08–.55)	.00 (.00–.30)	
WJ-III Quantitative Concepts	.69 (.18–.94)	.23 (.00–.72)	.26 (.00–.69)
	E ₁	E ₂	E ₃
SB SAS	.40 (.33–.48)		
Decoding score	.06 (.00–.16)	.00 (.00–.19)	
WJ-III Quantitative Concepts	.03 (.00–.26)	.00 (.00–.42)	.00 (.00–.40)
	A ₁	A ₂	A ₃
SB SAS	.73 (.56–.89)		
Decoding score	.63 (.45–.82)	.69 (.53–.77)	
WRAT	.62 (.25–.93)	.25 (.00–.48)	.00 (.00–.55)
	C ₁	C ₂	C ₃
SB SAS	.55 (.24–.71)		
Decoding score	.36 (.06–.56)	.04 (.00–.30)	
WRAT	.23 (.00–.67)	.71 (.00–.84)	.00 (.00–.82)
	E ₁	E ₂	E ₃
SB SAS	.40 (.33–.48)		
Decoding score	.07 (.00–.16)	.00 (.00–.19)	
WRAT	.08 (.00–.30)	.00 (.00–.31)	.00 (.00–.31)

Note. The 95% confidence intervals are in parentheses. The 95% confidence intervals are in parentheses. The first set of biometric factors measure the genetic (A₁), shared environmental (C₁), and nonshared environmental (E₁) influences between all three measurement factors. The second set (A₂, C₂, and E₂) measures the genetic and environmental influences between reading and math, outside of that explained by the SB SAS factor.

Finally, the third set of biometric factors (A₃, C₃, and E₃) measures the unique influences on math alone, outside of the variance explained by both SB SAS and reading. SB SAS = Stanford-Binet Intelligence Scale; RAN = Rapid Automatized Naming; WJ-III = Woodcock-Johnson III Tests of Achievement; WRAT = Wide Range Achievement Test—3.

^aMean score of Word Identification and Passage Comprehension subtests of the Woodcock Reading Mastery Tests—Revised.