CHILD DEVELOPMENT

Child Development, May/June 2016, Volume 87, Number 3, Pages 929–943

Preschool Drawing and School Mathematics: The Nature of the Association

Margherita Malanchini
Goldsmiths University of London and Tomsk State University

Maria G. Tosto
Tomsk State University and University of York

Victoria Garfield
University College London

Aysegul Dirik
Goldsmiths University of London and Unit for Social and Community Psychiatry (WHO Collaborating Centre for Mental Health Services Development), Queen Mary University of London

Adrian Czerwik
Goldsmiths University of London

Rosalind Arden
King's College London

Sergey Malykh
Psychological Institute of the Russian Academy of Education

Yulia Kovas
Goldsmiths University of London, Tomsk State University, King's College London and New York University in London

The study examined the etiology of individual differences in early drawing and of its longitudinal association with school mathematics. Participants (*N* = 14,760), members of the Twins Early Development Study, were assessed on their ability to draw a human figure, including number of features, symmetry, and proportionality. Human figure drawing was moderately stable across 6 months (average *r* = .40). Individual differences in drawing at age 4½ were influenced by genetic (.21), shared environmental (.30), and nonshared environmental (.49) factors. Drawing was related to later (age 12) mathematical ability (average *r* = .24). This association was explained by genetic and shared environmental factors that also influenced general intelligence. Some genetic factors, unrelated to intelligence, also contributed to individual differences in drawing.

Drawing can fulfill different functions in children's development, such as helping them to explore their ideas about the surrounding world, improve their spatial visualization and orientation skills, and enable them to create visual representations of their thoughts and feelings (Brook, 2009). Drawing skills emerge during the 2nd year of life and change significantly over the course of childhood (Braswell & Rosengren, 2008). Two-year-old children already show some understanding of the link between intention, action, and interpretation necessary for drawing production, and by the age of 3–4 children are able to apply this understanding to their drawings (Golomb, 1974).

A drawing can be described as having a dual nature: Not only is it a thing in itself (e.g., a mark on a page), but it also refers to a phenomenon in the internal or external world. To appreciate the dual nature of a picture, young children are required to flexibly adjust their thinking, for example, conceiving objects in two ways simultaneously

We gratefully acknowledge the ongoing contribution of the participants in the Twins Early Development Study (TEDS) and their families. TEDS is supported by a program grant (G0901245; and previously G0500079) from the U.K. Medical Research Council; our work on environments and academic achievement is also supported by grants from the U.S. National Institutes of Health (HD044454, HD046167, and HD059215). The research of Yulia Kovas and Sergey Malykh is supported by a grant from the Government of the Russian Federation (11.G34.31.0043). We are grateful to the students and the teachers of South Hampstead School, London who helped us with the development of the Drawing Maturity Scale.

Correspondence concerning this article should be addressed to Yulia Kovas, Department of Psychology, Goldsmiths University of London, Whitehead Building, New Cross, London, SE14 6NW UK. Electronic mail may be sent to y.kovas@gold.ac.uk.

DOI: 10.1111/cdev.12520

(Malchiodi, 2012). Cognitive flexibility, required for successful drawing, develops gradually, reaching appropriate levels around 4 years of age, when children start appreciating that pictures are representations of items as well as items themselves (Jolley, 2008). Drawing ability and cognitive flexibility continue to be positively related throughout childhood. For example, spatial drawing ability, the ability to use depth cues while reproducing a three-dimensional object, was found to be positively related (r = .35) to cognitive flexibility in 7- to 11-year-old children (Ebersbach & Hagedorn, 2011).

The notion that children's drawing ability is linked to their cognitive development has been around for more than a century. Cooke was the first to describe the successive developmental stages of children's drawing (Cooke, 1885, in Kamphaus & Pleiss, 1991). His work was followed by that of Ricci, who published his theory on children's drawing development in 1887 (Kamphaus & Pleiss, 1991). The scientific interest in children's drawing reached its peak at the beginning of the 20th century, with studies finding links between drawing and intelligence (Kamphaus & Pleiss, 1991). Piaget incorporated drawing in his developmental learning theory, describing it as an activity that indexed the child's cognitive maturity (Brook, 2009). Piaget described the developmental stages of drawing, arguing that drawing performance was emblematic of a child's cognitive competence. The stage-like development of the drawing ability is also exemplified by the observation that children can progress to a more complex representational ability only once they can master a more basic representation (e.g., Karmiloff-Smith, 1990).

Early research into human figure drawing ability resulted in the development of two measures that assess performance based on whether the necessary features of the human body are present: Goodenough–Harris Draw-a-Person (DAP) test (Harris, 1963) and the Draw-a-Child (DAC) scale (McCarthy, 1972). Both the DAP and the DAC are considered valid and reliable instruments for measuring the ability to draw a human figure, and have been shown to correlate with general intelligence in children (Reynolds, 1978). The tests are fast and easy to administer, and have been shown to be suitable for administration cross-culturally (Naglieri & Bardos, 1987). Both tests include clear scoring guidelines for clinicians and have been used in clinical settings, including with nonverbal children (Kamphaus & Pleiss, 1991).

However, recently the validity of using drawing tests as measures of cognitive ability and development has been challenged as only moderate relations are observed between drawing and cognitive abilities (e.g., Willcock, Imuta, & Hayne, 2011). For example, one study found a moderate positive correlation (.40) between performance in the DAP test and measures of cognitive ability, such as the Wechsler Intelligence Scale in 5- to 6-year-old children (Willcock et al., 2011). In addition, very little is known about the stability of the observed drawing–intelligence relation.

Moreover, the etiology of individual differences in human figure drawing and of its association with general cognitive ability (g) remains poorly understood. Drawing ability is characterized by a consistent, universal, and sequential progression, with very little influence from adults on drawing development (Kellogg, 1969, in Brook, 2009)—suggesting a genetically governed developmental process. However, individual differences in maturity of drawing are likely to be related to both genetic and environmental factors. For example, drawing quality and timing of progression to a more advanced stage of the drawing ability have been shown to be at least partly related to parental involvement (Dunst & Gorman, 2009). To date only one study has looked at the etiology of early human figure drawing ability and its relation with later g (Arden, Trzaskowski, Garfield, & Plomin, 2014). The study used data from the large longitudinal population-based Twins Early Development Study (TEDS) and found human figure drawing ability to be moderately heritable (approximately .3) with environmental factors being largely of the nonshared type (approximately .5). No sex differences were found in the etiology of the individual differences in human figure drawing ability: The same genetic and environmental factors were involved in differences among boys and differences among girls (Arden et al., 2014). The study also found that the relation between early drawing and g at age 14 (r = .20) was 99% explained by genetic factors.

It is possible that early human figure drawing ability may be particularly related to specific aspects of cognitive development. Mathematical ability may be strongly related to drawing, both because it is linked to general cognitive ability and because of several specific features that mathematics may share with human figure drawing. For example, awareness of number of body features, proportionality, appropriate use of space, and symmetry may all be specifically related to mathematical development.

As spatial ability has been found to be uniquely associated with mathematical ability (Rohde &

Thompson, 2007), the spatial nature of drawing may mean that drawing and mathematics are related above and beyond their relation with *g*. A large body of research has investigated the relation between mathematics and spatial ability—the ability to produce, recall, store, and modify images of objects (Lohman, 1996). Moderate correlations are consistently observed between mathematics and spatial ability across several measures, and this relation appears to be universal; it has been observed in different cultures (e.g., Krajewski & Schneider, 2009; van Garderen, 2006; Wei, Yuan, Chen, & Zhou, 2011). The association has been observed at every level of proficiency; for example, one study found that mathematically gifted adolescents also excelled in spatial ability (Dark & Benbow, 1991). Furthermore, the observed overlap between mathematics and spatial ability cannot be fully explained by their relation with *g*. One study of young adults found that spatial ability predicted mathematics achievement ($R^2 = .13$) after accounting for *g* (Rohde & Thompson, 2007). Support for the relation between mathematics and spatial ability was also observed at the etiological level (Tosto et al., 2014). Although both spatial and mathematical abilities at age 12 were only moderately heritable (.27 and .43, respectively), the correlation between them ($r = .43$) was largely explained by shared genetic factors (60%), most of which were also associated with *g*.

Existing measures of human figure drawing ability, such as the McCarthy scale (McCarthy, 1972), mainly focus on whether children's drawings contain the necessary bodily features (e.g., one head, two arms, two legs, etc.). Although valid and reliable, these measures might not be tapping into all of the cognitive skills involved in human figure drawing production. In fact, the relation between early drawing and several other cognitive skills, such as spatial and mathematical abilities, might be revealed by other characteristics of children's human figure drawing, including symmetry, proportionality, and the position of the drawing on the page. In order to overcome these limitations, we developed the Drawing Maturity Scale (DMS). The DMS explores eight aspects of early human figure drawing ability designed to tap into: (a) general cognitive development (emotionality, maturity of the lines, realism, and developmental stage of the drawing) and (b) spatial cognition (symmetry, proportionality, position of the drawing on the page, and percentage of the page that children used to produce their drawing). We hypothesize that both general and spatial features would be important for later mathematical development.

It is important to explore sex differences in early human figure drawing ability as they may be related to sex differences in mathematical ability. Previous research found mixed results in other cognitive domains, such as spatial and mathematical abilities. For most spatial and mathematical abilities no significant sex differences have been consistently found (Miller & Halpern, 2014). Some studies found a small female advantage in some aspects of the spatial domain, such as remembering object locations (e.g., Voyer, Potsma, Brake, & Imperato-McGinley, 2007). Males were found to show an advantage in other spatial tasks, such as two-dimensional and three-dimensional mental rotation (e.g., Voyer, Voyer, & Bryden, 1995). As these latter abilities seem more closely related to drawing, we predict a small male advantage in the human figure drawing task.

The present study used the data on drawing ability at ages 4 and 4½ and mathematical ability at age 12 from children participating in the longitudinal TEDS study. Measures of *g* at ages 4 and 12 were also available. We extended previous research on preschool drawing (Arden et al., 2014) by examining human figure drawing at age 4½, approximately 6 months after the initial assessment, and by developing a new measure of human figure drawing ability, the DMS. The goal of this research was threefold as we investigated: (a) stability of human figure drawing over a 6-month period in preschool boys and girls, (b) genetic and environmental etiology of human figure drawing ability at age 4½, and (c) the etiology of the association between early drawing of the human figure and later mathematical abilities accounting for general intelligence. In addition, the large sample used in this study allowed us to investigate sex differences in early human figure drawing ability.

Method

Sample

The sample of the present investigation includes twins from TEDS. All families with live twin births born in England and Wales between 1994 and 1996 were contacted by the Office of National Statistics on behalf of the study and over 15,000 families participated at first contact. Regular comparisons with the general population show that families in TEDS remain closely representative of the British population in socioeconomic distribution, ethnicity, and parental occupation (Oliver & Plomin, 2007). Informed, written consent was obtained from all of the families who agreed to take part in the study. TEDS focuses on investigating cognitive and behavioral traits across development. In 2014, all participants turned 18 and more than 8,000 twins remain actively involved in the study. More than 300 scientific papers have been published based on the TEDS data, and several spin-off projects have emerged from TEDS (for detailed information on TEDS, see Haworth, Davis, & Plomin, 2013). The number of participants, whose data were used in the present analyses, is reported separately for each measure in the Measures and Results sections.

Measures and Procedure

Human Figure Drawing Ability at Age 4

The Draw-a-Man test (McCarthy, 1972) was administered separately to each twin (*N*s = 14,580–4,999 monozygotic [MZ] and 9,581 dizygotic [DZ] twins) by their parents when they were 4 years old. Each child completed the drawing in the booklet provided to the family (one booklet per twin). All children were given the same instructions: "Draw me a picture of a girl (or boy if the child was a male). Do the best that you can. Make sure that you draw all of her." Parents were instructed to encourage the children in case they hesitated by saying things like, "You draw it all on your own, and I'll watch you. Draw the picture any way you like, just do the best picture you can." Parents were asked not to help the children by, for example, mentioning any missing body parts, and to make sure the child was finished with the drawing before putting the booklet away. The drawings were scored following the McCarthy's standardized procedure (McCarthy, 1972) by trained raters. These scoring criteria are based on the presence and absence of certain features (body parts such as head, hair, trunk, and arms) and require the rater to score each feature from 0 (*if the feature is absent*) to 1 (*if the feature is depicted well*). This created a maximum score of 12 for drawing ability at age 4.

Human Figure Drawing Ability at Age 4½

When the twins were 4½, they completed a similar task—the DAC test (McCarthy, 1972). This test was administered directly by experimenters during a home visit to a subset of twins (*N*s = 1,517–559 MZ and 958 DZ). The instructions, scoring, and coding procedure were the same as for the Draw-a-Man task (McCarthy, 1972). The scores ranged from 0 to a maximum 16 for drawing ability at age 4½.

In addition, the same drawings from a subsample of twins (*N*s = 517–163 MZ and 354 DZ) were evaluated using the DMS (see Appendix S2) that was specifically created and validated for this study. DMS aims to complement the existing measures of human figure drawing ability that mainly focus on the presence or absence of the main body parts. Other aspects of children's drawings, such as emotionality, position on a page, and symmetry, may also index the level of child's cognitive competence. DMS was specifically developed for this study to evaluate such features. The development of the DMS involved three piloting phases and the present study represents a further step toward the validation of the measure (see Appendix S1 for details of the scale development and use). The DMS includes eight items assessing emotionality, symmetry, maturity of lines, realism, proportionality, percentage of paper used, position of the drawing on the page, and the developmental stage shown by the picture. The scale has good internal validity, with a Cronbach's alpha of .86.

Mathematical Ability at Age 12

Teachers assessed twins' mathematical ability on the basis of the UK National Curriculum for Key Stage 3 (Qualifications and Curriculum Authority, 2003). Completed teacher questionnaires were received in the post. Teachers rated each child (*N*s = 6,140–2,245 MZ and 3,895 DZ) on a 9-point scale, from 1 being *very poor* to 9 being *exceptional performance*. The scale corresponds to the National Curriculum levels of achievement that UK teachers use for assessing pupils. For example, most 12-year-old students are expected to achieve Level 4 in mathematics. Teachers assessed four aspects of mathematical ability: using and applying mathematics; numbers; shapes, space, and measures; and handling data. Due to the high correlations between them (average $r = .80$; Oliver et al., 2004), the four aspects were collapsed into a composite score (teacher-rated mathematics composite at age 12).

At age 12, mathematical ability was also assessed by means of an Internet-based test battery, developed from the National Foundation for Educational Research (5–14) Mathematics Series (nFerNelson, 1999). The battery contained three subcomponents: understanding numbers, non-numerical processes and computation, and knowledge. Examples of questions include: (a) "Identify the missing number and type it into the appropriate box: $15 \times 6 = 90$; __ $\times 6 = 96$" and (b) the twins (*N*s = 8,846–3,098 MZ and 5,348 DZ) were presented with a picture of a partly shaded segmented circle and were asked to identify which fraction corresponded to the shaded part of the five options. Due to the strong correlations between the three tests, the three subcomponents were combined into a web-measured mathematics score at age 12 (for a detailed description of the measures, see Kovas, Haworth, Petrill, & Plomin, 2007; for additional information on the reliability and validity of this measure, see Haworth et al., 2007).

We included both teacher-rated and web-measured mathematics at age 12 as they reflect, at least partly, different aspects of mathematical ability. Of relevance to drawing, a more global teacher rating might reflect such specific characteristics as motivation and creativity.

General Cognitive Ability (g) at Ages 4 and 12

General cognitive ability was derived using principal component analysis, separately at each age. At age 4, parents administered two nonverbal and one verbal test to their children (*N*s = 15,123–5,209 MZ and 9,914 DZ). The nonverbal tests included an oddity task asking the twins to select a matching pair out of three items (Bayley, 1993) and puzzles (Raven, Court, & Raven, 1996). The verbal cognitive test assessed vocabulary (Receptive Picture Vocabulary Test, based on the Peabody Picture Vocabulary Test–Revised; Dunn & Dunn, 1981). Scores from these three tests were combined with scores of the Parent Report of Children's Abilities (PARCA; Saudino et al., 1998). Parents reported on their children's conceptual knowledge, syntax, expressive vocabulary, and abstract language. The PARCA shows good internal consistency ($\alpha = .74$) and moderate to strong correlation ($r = .55$) with the Mental Development Index of the Bayley Scales of Infant Development, 2nd ed. (Oliver et al., 2002).

At age 12, *g* was measured (*N*s = 7,280–2,696 MZ and 4,584 DZ) via two verbal tests: the multiple-choice information and multiple-choice vocabulary test (Wechsler Intelligence Scale for Children, 3rd ed. [WISC–III]-PI; Wechsler, 1992); and two nonverbal reasoning tests: the Picture Completion (WISC–III-UK; Wechsler, 1992) and Raven's Standard and Advanced Progressive Matrices (Raven, Raven, & Court, 1998).

Results

Descriptive statistics for all variables are presented in Table 1. All variables met the criteria for normal distribution. Only one twin out of each pair was randomly selected for phenotypic analyses in order to account for nonindependence of observation (i.e., the fact that the children are twins).

Gender Differences in Drawing Ability

Three univariate between-subjects analyses of variance (ANOVAs) were carried out on raw scores of one twin from each pair to identify potential sex differences in performance across the three drawing variables (Draw 4, Draw 4½, and DMS 4½). We found that girls outperformed boys on all measures of human figure drawing (see Table S4 in Appendix S3). Sex differences were significant ($p < .001$), with sex explaining between 4% and 7% of the variance in drawing ability. Levene's test showed that variances were comparable.

Despite these average sex differences, we did not expect sex differences in the etiology of individual differences in human figure drawing. Differences between groups may stem from different factors from those affecting individual differences within groups (see Kovas, Haworth, Dale, & Plomin, 2007). Previous research, applying a sex limitation model fitting to the human figure drawing data at age 4 found that the same genetic and environmental factors contributed to variation in drawing in boys and girls to the same extent (Arden et al., 2014).

Phenotypic Associations

Correlations between measures of drawing, mathematics, and *g* are shown in Table 2. Although the two drawing measures at age 4½ assessed different aspects of human figure drawing ability (number of features vs. drawing maturity), the correlation between them was strong (.74). The correlations between Draw 4 and Draw 4½, and between Draw 4 and DMS 4½ were .35 and .50, respectively —suggesting moderate stability of drawing across approximately 6 months. The drawing measures correlated modestly with the measures of mathematics and *g* at ages 4 and 12.

Predicting Mathematical Performance From Human Figure Drawing Maturity

Two separate linear regressions were carried out to examine the predictive power of human figure drawing at age 4½, as measured by the DMS, specifically developed for the purpose of this study to capture mathematically relevant processes. We evaluated the prediction from drawing to: (a)

Table 1
Descriptive Statistics for All Standardized Variables and Raw Scores for the Drawing Measures

Variable	*M*	*SD*	Min	Max	Skewness	Kurtosis	*N*[a]
Draw 4	0.00	1.01	−2.47	1.89	−.65	.006	7,290
Draw 4 raw score	6.71	2.88	0	12	−.70	.032	7,290
Draw 4½	−0.01	1.01	−1.88	2.35	−.33	−.50	758
Draw 4½ raw score	6.89	3.91	0	16	−.31	−.62	758
DMS	−0.06	0.99	−2.38	2.05	−.22	−.63	258
DMS raw score	24.17	6.67	9.33	38.00	−.27	−.63	258
Teacher Math 12	−0.02	0.93	−2.67	2.89	−.02	.22	3,070
Web Math 12	0.03	0.95	−2.96	1.83	−.76	.10	4,423
g 4	0.02	0.96	−3.00	2.66	−.40	−.89	7,561
g 12	0.01	0.98	−2.86	2.76	−.37	−.19	3,640

Note. The variables were standardized on the whole sample and then one twin out of each pair was selected for further analysis in order to control for nonindependence of observation. The means and standard deviations were estimated after excluding outliers. Deviations of means and standard deviations from 0 and 1, respectively, are a result of this selection. Draw = drawing score; DMS = Drawing Maturity Scale score; Teacher Math 12 = teacher ratings of mathematics at age 12; Web Math 12 = mathematics web test total score age 12; *g* = general cognitive ability.
[a]One twin out of each air was selected to account for nonindependence of observation. We utilized all data available for each measure. The small *N*s for the 4½ drawing measure and for the DMS measure are based on the in-home data collection wave that involved only a subsample of the Twins Early Development Study.

Table 2
Phenotypic Correlations and Number of Participants (N)

	Draw 4	Draw 4½	DMS 4½	MT 12	MW 12	*g* 4	*g* 12
Draw 4	1						
N							
Draw 4½	.35**	1					
N	732						
DMS 4½	.50**	.74**	1				
N	251	258					
MT 12	.20**	.14**	.31**	1			
N	2,960	339	141				
MW 12	.20**	.20**	.15**	.52**	1		
N	4,074	462	164	2,412			
g 4	.28**	.28**	.34**	.21**	.23**	1	
N	7,302	759	258	3,062	4,205		
g 12	.18**	.26**	.24**	.47**	.63**	.25**	1
N	3,512	410	145	2,145	3,458	3,627	

Note. Draw 4 = drawing score at age 4; Draw 4½ = drawing score at age 4½; DMS 4½ = Drawing Maturity Scale scores; *g* 4 and 12 = general cognitive ability at ages 4 and 12; MT 12 = teacher-rated mathematics at age 12; MW 12 = web-measured mathematics at age 12; *N* = only one twin out of each pair was randomly selected.
*$p < .05$. **$p < .01$ (two tailed).

teacher-rated mathematics at age 12 and (b) web-measured mathematics at age 12. DMS at age 4½ was a significant predictor of the two mathematical outcomes, explaining 12.3% and 4.7% of the variance in teacher-rated and web-measured mathematics at age 12, respectively.

The regressions were repeated, including a contemporaneous measure of *g* (at age 12) to assess whether DMS at age 4½ had any specific association with later mathematics beyond a more general association with *g*. The results, reported in Table S4 (Appendix S3), show that drawing ability at age 4½ measured with the DMS remained a significant predictor of teacher-rated mathematics at age 12 after accounting for *g* at the same age. The overall model was significant, $F(2, 105) = 21.312$, $p < .001$, $R^2 = .29$. In contrast, general cognitive ability at age 12 was the only significant predictor of web-assessed mathematics at age 12, $F(2, 141) = 46.041$, $p < .001$, $R^2 = .39$.

The same regressions were run replacing the DMS with the McCarthy scale at ages 4 and 4½.

The McCarty scale at age 4 was a significant predictor of the two measures of mathematics at age 12, explaining 3.8% of the variance in teacher-rated mathematics and 4.5% of the variance in web-measured mathematics. When *g* was included in the analyses, the McCarthy measure at age 4 remained a significant predictor of teacher-rated and web-assessed mathematics at age 12, although *g* explained most of the variance. The McCarthy scale at age 4½ was a significant predictor of the two mathematical outcomes, explaining 1.9% and 4.4% of the variance in teacher-rated and web-measured mathematics at age 12, respectively. After controlling for *g*, the McCarthy measure at age 4½ was not a significant predictor of later mathematical ability. The results suggest that most of the longitudinal association between early human figure drawing measures and later mathematical ability is not unique to mathematics, but rather reflects the stability of general cognitive ability. The DMS scale, developed specifically to tap into mathematically relevant abilities, was largely not uniquely related to mathematics at age 12, as most of the variance in the association between drawing at age 4½ and mathematics at age 12 was also shared with *g* at age 12.

Genetic and Environmental Etiology of Individual Differences in Human Figure Drawing Ability at Age 4½

Comparing similarities between MZ and DZ twins allows for an estimation of the relative contribution of genetic and environmental factors to individual differences in a given trait (Rutter, 2006). The ACE model allows us to assess the proportion of the variance in a phenotypic trait that is explained by additive genetic (A), shared environment (C), and nonshared environment (E) by comparing the similarity between MZ twins, who share 100% of their genes, and DZ twins, who on average share 50% of their segregating genes. Consequently, the ACE model posits that similarities between MZ twins for a specific trait could be explained by shared genetic and/or common environmental factors, whereas differences between MZ twins are due to nonshared environmental factors and measurement error. On the other hand, differences between DZ twins could be due not only to nonshared environmental influences and measurement error, but also to their genetic differences (Rijsdijk & Sham, 2002). Genetic influence can be estimated by comparing intraclass correlations for MZ and DZ twins. A greater similarity between MZ twins than between DZ twins for a specific trait indicates a degree of genetic influence on the variance of that specific trait. Heritability, the amount of variance in a trait that can be attributed to genetic variance, can be calculated as double the difference between the MZ and DZ twin correlations.

The univariate ACE model-fitting analysis is a more comprehensive way of estimating the proportion of phenotypic variance that can be attributed to genetic and environmental factors. As opposed to the estimates that can be derived from cross-twin correlations, model fitting allows us to assess the goodness of fit of the model including the latent variables A, C, and E, by comparing it to the saturated model (which is the model based on the observed data), and to more parsimonious models (e.g., models only including the latent factors A and E, or A and C, or only the latent factor E). Models are usually compared using maximum likelihood, Akaike's information criterion, or Bayesian information criterion. The univariate model also estimates confidence intervals for all parameters (see Plomin, DeFries, Knopik, & Neiderhiser, 2013, for details of the twin methodology; see Neale, Boker, Bergeman, & Maes, 2005, for the model-fitting procedures). We conducted the univariate ACE model-fitting analysis to assess the etiology of individual differences separately in each measure. The results of these analyses are presented in Table S6, with the exception of the results for human figure drawing ability (McCarthy measure) at age 4½, which are presented below.

As can be seen in Table 3, MZ correlations exceeded DZ correlations—indicating significant genetic influences on this ability. As the DZ correlations were more than half of the MZ correlations, significant shared environmental effects were also indicated. As MZ correlations were only about .5, about half of the variance in drawing was explained by nonshared (individual-specific) factors, which also include measurement error.

The model-fitting analysis showed that genetic factors explained 21% of the variance in drawing ability at age 4½. Shared and nonshared environment explained 30% and 49% of the variance, respectively. These results are highly similar to previous results from the TEDS sample at age 4 (Arden et al., 2014).

Etiology of the Relation Between Early Human Figure Drawing Ability and School Mathematics at Age 12

The univariate method can be extended to assess the etiology of the covariation between variables. We used the trivariate Cholesky decomposition

Table 3
Intraclass Correlations, ACE Estimates, and Fit Indices for Drawing at Age 4½

	*r*MZ (95% CI)		*r*DZ (95% CI)
Drawing age 4½	.52 (.43–.60)		.38 (.30–.46)
N	279		478
	A	C	E
ACE estimates			
	.21 (.09–.41)	.30 (.30–.45)	.49 (.21–.58)
	−2LL	AIC	BIC
Goodness-of-fit indices for drawing at age 4½ (McCarthy measure)			
Saturate model	4,138.84	1,112.84	−9,379.98
ACE model	4,144.92	1,106.92	**−9,427.51**

Note. Drawing age 4½ = as measured by McCarthy Scale; *r*MZ = intraclass correlation for monozygotic twins; *r*DZ = intraclass correlation for dizygotic twins; 95% CI = 95% confidence intervals; AIC = Akaike's information criterion. Both same-sex and opposite-sex DZ twin pairs were included in the analyses. A smaller, more negative, Bayesian information criterion (BIC in bold) indicates better fit.

model (Neale et al., 2005) to examine to what extent common genetic and environmental influences explained the correlations between our three variables of interest: drawing, mathematics, and *g*. The Cholesky model decomposes phenotypic variance and covariance between traits into common and independent genetic (A), shared environmental (C), and nonshared environmental (E) sources of variance and covariance (e.g., Wang et al., 2014). The model works similarly to a hierarchical regression analysis, as the independent contribution of a predictor variable to the dependent variable is estimated after accounting for the variance it shares with other predictors (Luo, Kovas, Haworth, Dale, & Plomin, 2011).

Table 4 reports cross-twin–cross-trait correlations for each pair of variables. Cross-twin–cross-trait correlations describe the association between two variables, with Twin 1 score on Variable 1 correlated with Twin 2 score on Variable 2. Cross-twin–cross-trait correlations were calculated separately for MZ and DZ twins. A higher cross-twin–cross-trait correlation for MZ than for DZ twins indicates that genetic factors have a degree of influence on the phenotypic relation between the two traits. For example, the correlations between *g* for Twin 1 and drawing for Twin 2 are .28 for MZ and .19 for DZ twins. The correlations were very similar when the traits were swapped for Twin 1 and Twin 2.

Table 4
Cross-Twin–Cross-Trait Correlations for the Associations Between g *at Age 4, Drawing at Age 4, and Mathematics (Both Teacher Rated and Web Measured) at Age 12, and for the Associations Between the McCarthy Measure of Drawing at Age 4½,* g *at Age 4, and Mathematics at Age 12*

Pairs of variables	*r*MZ	*r*DZ
g 4 and Drawing 4	.28	.19
g 4 and Math T12	.16	.14
Drawing 4 and Math T12	.22	.13
g 4 and Math W12	.23	.18
Drawing 4 and Math W12	.21	.15
g 4 and Drawing 4½	.23	.18
Drawing 4½ and Math T12	.31	.09
Drawing 4½ and Math W12	.27	.1

Note. g = general cognitive ability at age 4; Drawing 4 = drawing score at age 4 measured with the McCarthy Draw-a-Man scale; Math T12 = teacher ratings of mathematics at age 12; Math W12 = mathematics web test scores at age 12; Drawing 4½ = drawing score at age 4½ measured with the McCarthy Draw-a-Child scale; *r*MZ = intraclass correlation for monozygotic twins; *r*DZ = intraclass correlation for dizygotic twins.

Four separate trivariate models were run using the McCarthy drawing measure at ages 4 and 4½ and mathematics measures (teacher rated and web assessed) at age 12, with *g* at age 4 added to each model. The results of the analyses of the two drawing measures were similar overall, but due to a sample size reduction, the measure that children completed at age 4½ produced very wide confidence intervals. We therefore report only the two multivariate analyses run on drawing at age 4, for which the largest sample size was available.

The first model considered the trivariate association between *g* at age 4, human figure drawing at age 4, and teacher-rated mathematics at age 12. The best model to fit the data was selected on the basis of goodness of fit, as previously done for the univariate analysis. The full ACE model was found to be the best fit for the data (see Appendix S4, Table S7 for goodness of fit and standardized squared path estimates).

The results of the trivariate Cholesky decomposition, shown in Figure 1, indicate that the etiology of *g* at age 4 was due largely to shared environmental factors (C1 = 62%), with moderate genetic influence (A1 = 24%) and a smaller portion of variance explained by nonshared environmental factors (E1 = 14%) that also include error. About 5% of the genetic factors (the path from A1 to drawing) and 4% of the shared environmental factors (the path from C1 to drawing) that influenced *g* also influenced individual differences in drawing at age 4.

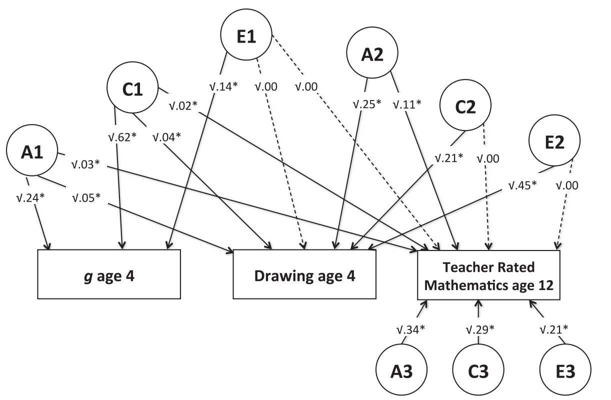

Figure 1. Trivariate Cholesky decomposition model for the association between *g* at age 4, drawing at age 4, and teacher-rated mathematics at age 12. A1 = additive genetic effects common to Variable 1 (*g* at age 4), Variable 2 (drawing at age 4), and Variable 3 (mathematics); A2 = additive genetic effects common to Variables 2 and 3; A3 = additive genetic effects specific to Variable 3; C1 = shared environmental effects common to Variables 1, 2, and 3; C2 = shared environmental effects common to Variables 2 and 3; C3 = shared environmental effects specific to Variable 3; E1 = nonshared environmental effects common to Variables 1, 2, and 3; E2 = nonshared environmental effects common to Variables 2 and 3; E3 = nonshared environmental effects specific to Variable 3. The dashed lines represent nonsignificant paths. *indicates significant estimates.

Similarly, 3% and 2% of genetic and shared environmental factors, respectively, influenced both *g* at age 4 and teacher-rated mathematics at age 12.

Variation in mathematics at age 12 was explained by genetic (48%), shared environmental (31%), and nonshared environmental (21%) factors. For example, looking at the model in Figure 1, the heritability of teacher-rated mathematics at age 12 can be obtained by adding up the estimates of all the paths linking the latent factors A to teacher-rated mathematics at age 12 (A1, A2, and A3 in this model). Therefore, the heritability of teacher-rated mathematics at age 12 is estimated at .48 ($\sqrt{.34} + \sqrt{.11} + \sqrt{.03} = \sqrt{.48}$). The results also show that etiology of the individual differences in mathematics at age 12 is largely independent from that of general cognitive ability and drawing ability at age 4. The strongest influences on mathematics come from the latent factors A3 (34%), C3 (29%), and E3 (21%), which represent the proportion of genetic (A3), shared environmental (C3), and nonshared environmental (E3) variance in teacher-rated mathematics at age 12 that is not shared with *g* and drawing at age 4. The latent factors A2, C2, and E2 indicate the proportion of the variance in the etiology of mathematics at age 12 that is shared with drawing at age 4, independent of *g* at age 4. As indicated by the path from latent variable A2 to mathematics at age 12, 11% of the heritability of teacher-rated mathematics at age 12 was shared with drawing, but independent of *g* at age 4. As indicated by the path from the latent variable A1 to mathematics at age 12, only 3% of the heritability of mathematics at age 12 was shared with both drawing and *g* at age 4. Shared and nonshared environmental effects were largely specific to each trait. For example, the paths from latent variables E1 to drawing and mathematics, and from E2 to mathematics were nonsignificant.

Another way of looking at the etiology of the interrelation across the measures is to look at the genetic, shared environmental, and nonshared environmental correlations between variables. From these estimates it is possible to derive the proportion of the phenotypic correlation between variables that can be attributed to genetic, shared, and nonshared environmental influences. The proportion of the phenotypic correlation that can be attributed to genetic influences common to both variables is known as bivariate heritability. In the same way, bivariate shared environment is the proportion of the phenotypic correlation that can be attributed to shared environmental factors common to both variables. Finally, bivariate nonshared environment is the proportion of the phenotypic correlation that can be attributed to nonshared environmental factors common to both variables. Table 5 presents the pairwise phenotypic correlations among the three measures, which were overall modest (average .24). The table also presents genetic and environmental correlations. Average genetic correlation was moderate (.40) and average shared environmental correlation was modest (.28). The average nonshared environmental correlation was negligible (.04).

The bivariate heritability and environmentalities for each pairwise association were derived using the following formula: $(\sqrt{h^2}\ (\text{draw}) \times \sqrt{h^2}\ (\text{math}) \times r_G)/r_P$ for genetic effects, $(\sqrt{c^2}\ (\text{draw}) \times \sqrt{c^2}\ (\text{math}) \times r_C)/r_P$ for shared environment, and $(\sqrt{e^2}\ (\text{draw}) \times \sqrt{e^2}\ (\text{math}) \times r_E)/r_P$ for nonshared environment (see Table 5). The largest proportion of the modest phenotypic correlation (.24) between drawing at age 4 and teacher-rated mathematics at age 12 was explained by genetic factors (82%), with a smaller portion explained by shared environment (18%). The modest phenotypic association (.21) between *g* at age 4 and teacher-rated mathematics at age 12 was explained mostly by shared environment (56%) and genetic factors (40%), with a minor proportion explained by nonshared environmental influences (5%). The modest correlation (.28) between drawing

Table 5

Phenotypic (r_P)*, Genetic* (r_G)*, Shared* (r_C)*, and Nonshared Environmental* (r_E) *Correlations and Bivariate Heritability/Environmentalities for the Trivariate Association Between* g *at Age 4, Drawing at Age 4, and Teacher-Rated Mathematics at Age 12*

Variables	Bivariate h^{2a} and r_G (95% CI)	Bivariate c^2 and r_C (95% CI)	Bivariate e^2 and r_E (95% CI)
g 4 and Drawing	.38	.56	.06
r_P = .28 (.27–.30)	41 (.30 to .53)	.40 (.33 to .47)	.07 (.03 to .10)
g 4 and Math T12	.40	.56	.04
r_P = .21 (.18–.24)	.25 (.12 to .37)	.27 (.17 to .37)	.06 (.04 to .13)
Drawing 4 and Math T12	.82	.18	.00
r_P = .24 (.21–.26)	.53 (.37 to .70)	.16 (−.02 to .34)	−.01 (−.08 to .04)

Note. g = general cognitive ability at age 4; Drawing = drawing ability at age 4 measured with the McCarthy scale; Math T12 = teacher-rated mathematics at age 12. The phenotypic correlation estimates are slightly different from the correlations previously reported as these were obtained after the data were regressed for age and gender, as is standard practice in the data preparation for ACE model fitting.

[a]Bivariate heritability/environmentalities = the proportion of the phenotypic correlation (r_P) explained by common genetic, shared, and nonshared environmental factors.

and *g* at age 4 was explained mostly by shared environmental (56%) and genetic (38%) influences.

Another model, out of the four Cholesky decompositions, examined the association between *g* at age 4, human figure drawing at age 4, and web-assessed mathematics at age 12 (see Appendix S5). Overall, results were consistent with those obtained for teacher-rated mathematics at age 12. Individual differences in general cognitive ability at age 4, drawing ability at age 4, and web-assessed mathematics at age 12 were found to have largely different etiology. Approximately 5% of the heritability of drawing at age 4 and 6% of the heritability of web-measured mathematics at age 12 were due to genetic factors shared with *g* at age 4. Drawing at age 4 and web-measured mathematics at age 12 shared only 3% of their heritability beyond that already shared with *g* at age 4.

In addition to the four analyses, another trivariate Cholesky decomposition was run looking at the etiology of the covariation between human figure drawing at age 4, *g* at age 12, and teacher-rated mathematics at age 12 (see Appendix S6). Results were consistent with those obtained when looking at the covariance between *g* at age 4, human figure drawing at age 4, and teacher-rated mathematics at age 12. Approximately 6% of the heritability of *g* at age 12 was due to genetic factors shared with drawing at age 4. Interestingly, around 12% of the heritability of teacher-rated mathematics at age 12 was due to genetic factors shared with drawing at age 4 that were not shared with *g* at age 12. Independent of drawing at age 4, *g* at age 12 and teacher-rated mathematics at age 12 shared 6% of their heritability.

Discussion

The aim of the present study was to investigate early human figure drawing ability in its relation to later mathematical ability. Specifically, we addressed three main questions: (a) the stability of drawing over 6 months in preschool boys and girls, (b) the etiology of individual differences in human figure drawing ability at age 4½, and (c) the etiology of the longitudinal relation between preschool drawing, school mathematics, and *g* (measured at ages 4 and 12).

Human figure drawing ability was found to be reasonably stable from 4 to 4½ years of age (average r = .42). Human figure drawing ability at age 4 was measured with the McCarthy Draw-a-Man scale (1972); human figure drawing ability at age 4½ was measured using the McCarthy DAC scale (1972), as well as the DMS developed specifically for the purpose of this investigation. The relation between the two drawing measures reflects the stability of individual differences in drawing performance over development and also demonstrates the validity of the new DMS scale. Although 6 months seems a short period, it is a large portion of a child's life at this age, with several stages happening during this time in the development of drawing production (Malchiodi, 2012).

To our knowledge this study is the first to test the stability of human figure drawing ability longitudinally in a large representative sample. In fact, research findings on drawing development are too often limited by small sample sizes that lack adequate power to allow longitudinal analyses.

Furthermore, our sample was homogeneous in age (all children were close to ages 4 and 4½ when tested); this is often not the case in the drawing literature, which is characterized by large age ranges. Large age gaps between participants are particularly problematic if the aim is to assess drawing development during childhood, as drawing production changes significantly over a relatively short developmental time.

Our large sample also allowed us to explore sex differences in human figure drawing ability with sufficient power. We found that sex differences in drawing performance explained between 4% and 7% of the variance in human figure drawing ability at ages 4 and 4½, with girls scoring higher than boys at both ages and across the three measures (McCarthy Draw-a-Man, McCarthy DAC, and DMS). However, the etiology of individual differences was the same for boys and girls, as suggested by findings from a previous study that ran sex limitation models on the same drawing data at age 4 (Arden et al., 2014). This is consistent with a previous investigation of into the etiology of spatial ability and its relation with mathematics that found no gender differences in the etiology of both abilities (Tosto et al., 2014).

The strong correlation observed between the two measures of human figure drawing ability at age 4½ (r = .74) suggests that the newly developed DMS scale is a valid instrument to measure individual differences in drawing performance in preschool children. The observed association between the DMS and the McCarthy DAC scale at age 4½ partly reflects general cognitive ability. The association between drawing ability and g was found to be largely stable from age 4½ (average r = .30) to age 12 (average r = .25). This corroborates previous findings of a stable association between drawing at age 4 and g at age 14 (Arden et al., 2014) and extends it to another measure of human figure drawing ability (i.e., McCarthy measures as well as the DMS).

Similarly, the results on the etiology of individual differences in early human figure drawing ability at age 4½ were consistent with those previously obtained with a sample of 4-year-old twins (Arden et al., 2014): Genetic (21%) and shared environmental (30%) influences were modest and nonshared (child specific) environmental influences explained a larger portion (49%) of the variance in drawing ability at age 4½. However, a proportion of this variance could be due to error of measurement, as nonshared environmental effects include measurement error (Plomin, 2011), which is an important consideration when assessing young children's abilities.

A further main aim of this investigation was to explore the specificity of the drawing–mathematics association over development. When we examined the association at the phenotypic level, we found that most of the variance in the prediction from drawing at age 4½ to mathematical ability at age 12 was also shared with g. In fact, regression analyses showed that drawing at age 4½ measured using the McCarthy DAC scale did not remain a significant predictor of mathematical ability at age 12 (both teacher rated and web measured) after controlling for g at age 12. The prediction from the DMS at age 4½, developed specifically to tap into the drawing–mathematics association, to teacher-rated mathematics at age 12 remained significant after accounting for g; however, g explained a large portion of their relation. On the other hand, the prediction from the DMS to web-measured mathematics at age 12 was not significant beyond g at age 12. Our results suggest that the relation between human figure drawing and mathematics is not specific and in fact is mostly accounted for by general intelligence. These results are consistent with previous findings demonstrating links between drawing ability and other cognitive abilities (Gottling, 1990).

The overall absence of a unique relation between drawing and mathematics goes against our prediction. We developed the DMS to specifically tap into those aspects of human figure drawing that could be more closely related to spatial and mathematical development, such as the position of the drawing on the page and the proportionality of the drawing. In fact, as previously observed for spatial ability (e.g., Rohde & Thompson, 2007), we expected the relation between human figure drawing ability at age 4½ measured with the DMS and mathematical ability at age 12 to extend beyond their association with g. This was only partly supported when considering the relation between the DMS and teacher-rated mathematics at age 12.

The lack of specificity of the relation between the DMS and the two mathematics outcomes might also reflect the broad range of skills assessed by this new measure of human figure drawing ability. We explored this issue further by examining the prediction from factor 2 of the DMS (assessing the position of the drawing on the page and the percentage of paper occupied by the drawing), hypothesizing that these features could be more reflective of later mathematics competence. However, after controlling for g, factor 2 of the DMS did not remain a significant predictor of mathematics at age 12 (see

Table S5). It is possible that other aspects of drawing, not examined by the DMS, may be specifically associated with mathematics performance. On the other hand, it is also possible that human figure drawing ability is associated with other aspects of mathematics performance, such as numerical magnitude comparison and number line estimation.

At the etiological level, the observed associations between human figure drawing at age 4 and general cognitive ability contemporaneous to drawing and mathematics at age 12 were explained by overlapping genetic and shared environmental effects. The results of the trivariate analyses, where *g* at age 4 was also included, showed that common genetic influences on early drawing and later mathematics include mostly factors in common with *g*. The genetic associations between drawing at age 4 and mathematics at age 12 were largely similar after accounting for *g* with an indication of a slightly stronger relation for the teacher-rated measure.

Overall, the results are consistent with the "generalist genes" account of learning abilities and disabilities that proposes that most of the genes implicated in cognitive abilities and academic achievement are general as opposed to specific to each domain (Plomin & Kovas, 2005). The account is grounded in the two concepts of pleiotropy (one gene affects many traits) and polygenicity (several genes influence one trait) and proposes that genetic influences on different abilities, as well as disabilities, overlap. Several studies using multivariate genetic analyses have found support for the generalist genes theory (e.g., Kovas, Harlaar, Petrill, & Plomin, 2005; Plomin & Kovas, 2005). Additionally, studies using genome-wide complex trait analysis (GCTA; Yang et al., 2010), a method that allows for the estimation of heritability of complex traits from DNA samples of unrelated individuals, have also found support for the generalist genes account. The genetic correlations between *g* and language, reading, and mathematical abilities obtained using GCTA exceeded .70 (Trzaskowski et al., 2013); these results, consistent with findings from twin studies, indicate the pleiotropic effects of the genes implicated in the variation in cognitive abilities. Evidence in support for the generalist genes hypothesis also comes from molecular genetic research. For example, most of the single nucleotide polymorphisms associated with early reading ability were also found to be associated with aspects of mathematics and general cognitive ability as well as with other components of literacy (Haworth, Meaburn, Harlaar, & Plomin, 2007).

The links between human figure drawing, mathematics, and *g* could be related to motor development. For example, drawing scores may reflect maturity of lines that in turn depend on motor skills; close links have been found between the development of cognitive and motor skills in nonclinical populations (e.g., Martin, Tigera, Denckla, & Mahone, 2010). Moreover, recent neuroimaging research suggests that partly overlapping cortical and subcortical brain regions are associated with the development of both general cognitive ability and motor skills (Pangelinan, Zhang, VanMeter, Clark, & Hatfield, 2011).

A number of studies have explored the relation between cognitive and motor abilities in clinical populations (e.g., Davis, Pass, Finch, Dean, & Woodcock, 2009). In particular, children with spina bifida—a congenital neurodevelopmental disorder characterized by motor deficits—showed impairments in several aspects of mathematics performance, from counting to magnitude comparison (i.e., English, Barnes, Taylor, & Landry, 2009). It is possible that deficits in fine motor skills, including finger function and precision in upper limb control, in infants and toddlers with spina bifida place early constraints on those aspects of counting and simple arithmetic that are supported by pointing, touching, and finger counting. Further research is needed into the relation between drawing ability and mathematics in children characterized by restricted motor abilities over an extended developmental time.

A key strength of this study is its large longitudinal sample that allowed us to have sufficient statistical power to examine the stability of human figure drawing ability, gender differences in performance, the etiology of human figure drawing ability, and its longitudinal association (phenotypically and etiologically) with *g* and mathematics. A second strength of this investigation is that our data were obtained from several sources (parental reports, teacher assessments, child performance on cognitive and mathematical ability tests, and child drawing evaluations by trained raters)—minimizing the possible biases of single source information.

Although the richness of our twin sample allowed for the in-depth investigation of the drawing–mathematics relation, the fact that the children in this sample are twins comes with a few limitations. In fact, twin studies are based on a number of assumptions. One of these assumptions is the idea that environmental similarity is the same for MZ and DZ twin pairs growing up in the same family (equal environments assumption). Although evidence suggests that MZ twins are more likely to experience similar environments than DZ twins (e.g., they tend to be treated more similarly, more often share the same playmates, etc.), sharing more environmental experiences was not found to impact on the degree of their phenotypic concordance (Kendler, Kessler, Neale, Heath, & Eaves, 1993). A second limitation, particularly relevant to the present study, is the fact that evidence suggests that twins might be at a slight disadvantage during gestation and early development if compared to singletons (Martin, Boomsma, & Machin, 1997). Therefore, before these findings can be extended, replication in a general population is required. A further limitation of the present investigation was that the DMS ratings were only available on a relatively small subsample of the children. Further investigations with a larger sample is necessary to test the added value of this scale over the traditional "number of features" assessments of early drawing.

Overall, the present investigation of individual differences in preschool human figure drawing ability represents a step forward in our understanding of the mechanisms through which early drawing ability relates to overall cognitive development.

References

Arden, R., Trzaskowski, M., Garfield, V., & Plomin, R. (2014). Genes influence young children's human figure drawings and their association with intelligence a decade later. *Psychological Science*, *25*, 1843–1850. doi:10.1177/0956797614540686

Bayley, N. (1993). *Bayley Scales of Infant Development* (2nd ed.). San Antonio, TX: Psychological Corporation.

Braswell, G., & Rosengren, K. (2008). The interaction of biomedical and cognitive constraints in the production of children-s drawings. In C. Lange-Kuttner & A. Vinter. *Drawing and the nonverbal mind. A life-span perspective* (pp. 123–138). New York, NY: Cambridge University Press.

Brook, M. (2009). Drawing, visualisation and young children's exploration of "big ideas." *International Journal of Science Education*, *31*, 319–341. doi:10.1080/09500690802595771

Dark, V. J., & Benbow, C. P. (1991). Differential enhancement of working memory with mathematical versus verbal precocity. *Journal of Educational Psychology*, *83*, 48–60. doi:10.1037/0022-0663.83.1.48

Davis, A. S., Pass, L. A., Finch, W. H., Dean, R. S., & Woodcock, R. W. (2009). The canonical relationship between sensory-motor functioning and cognitive processing in children with attention-deficit/hyperactivity disorder. *Archives of Clinical Neuropsychology*, *24*, 273–286. doi:10.1093/arclin/acp032

Dunn, L., & Dunn, L. (1981). *The Peabody Picture Vocabulary Test–Revised*. Circle Pines, MN: American Guidance Service.

Dunst, C. J., & Gorman, E. (2009). Development of infant and toddler mark making and scribbling. *CELLreviews*, *2*(2), 1–16. Retrived from http://www.earlyliteracylearning.org/cellreviews/cellreviews_v2_n2.pdf

Ebersbach, M., & Hagedorn, H. (2011). The role of cognitive flexibility in the spatial representation of children's drawings. *Journal of Cognition and Development*, *12*, 32–55. doi:10.1080/15248372.2011.539526

English, L., Barnes, M. A., Taylor, H. B., & Landry, S. H. (2009). Mathematical development in spina bifida. *Developmental Disabilities Research Review*, *15*, 28–34. doi:10.1002/ddrr.48

Golomb, C. (1974). *Young children's sculpture and drawing*. Cambridge, MA: Harvard University Press.

Gottling, S. H. (1990). The validity of the "Draw-a-Person: A quantitative scoring system" for kindergarten children (Doctoral Dissertation, Ohio State University, 1990). *Dissertation Abstracts International*, *51*, 828.

Harris, D. B. (1963). *Children's drawings as measures of intellectual maturity*. New York, NY: Harcourt, Brace & World.

Haworth, C. M. A., Davis, O. S. P., & Plomin, R. (2013). Twins Early Development Study (TEDS): A genetically sensitive investigation of cognitive and behavioural development from childhood to young adulthood. *Twin Research and Human Genetics*, *16*, 117–125. doi:10.1017/thg.2012.91

Haworth, C. M., Harlaar, N., Kovas, Y., Davis, O. S., Oliver, B. R., Hayiou-Thomas, M. E., . . . Plomin, R. (2007). Internet cognitive testing of large samples needed in genetic research. *Twin Research and Human Genetics*, *10*, 554–563.

Haworth, C. M. A., Meaburn, E. L., Harlaar, N., & Plomin, R. (2007). Reading and generalist genes. *Mind, Brain, Education*, *1*, 173–180. doi:10.1111/j.1751-228X.2007.00018.x

Jolley, R. (2008). Children's understanding of the dual nature of pictures. In C. Lange-Kuttner & A. Vinter (Eds.), *Drawing and the nonverbal mind. A life-span perspective* (pp. 86–103). New York, NY: Cambridge University Press.

Kamphaus, R. W., & Pleiss, K. (1991). Draw-a-Person techniques: Tests in search of a construct. *Journal of School Psychology*, *29*, 395–401. doi:10.1016/0022-4405(91)90026-N

Karmiloff-Smith, A. (1990). Constraints on representational change: Evidence from children's drawing. *Cognition*, *34*, 57–83. doi:10.1016/0010-0277(90)90031-E

Kendler, K. S., Kessler, R. C., Neale, M. C., Heath, A. C., & Eaves, L. J. (1993). The prediction of major depression in women: Toward an integrated model. *American Journal of Psychiatry*, *150*, 1139–1148.

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

Kovas, Y., Haworth, C. M. A., Dale, P. S., & Plomin, R. (2007). The genetic and environmental origins of learning abilities and disabilities in the early school years. *Monographs of the Society for Research in Child Development*, *72*(3), doi:10.1111/j.1540-5834.2007.00453.x

Kovas, Y., Haworth, C. M. A., Petrill, S. A., & Plomin, R. (2007). Mathematical ability of 10-year-old boys and girls: Genetic and environmental etiology of normal and low performance. *Journal of Learning Disabilities*, *40*, 554–567.

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

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

Luo, Y. L. L., Kovas, Y., Haworth, C. M., Dale, P. S., & Plomin, R. (2011). The etiology of mathematical self-evaluation and mathematics achievement: Understanding the relationship using a cross-lagged twin study from ages 9 to 12. *Learning and Individual Differences*, *21*, 710–718.

Malchiodi, C. A. (2012). *Understanding children's drawings*. New York, NY: Guilford.

Martin, N., Boomsma, D., & Machin, G. (1997). A twin-pronged attack on complex traits. *Nature Genetics*, *17*, 387–392.

Martin, R., Tigera, C., Denckla, M. B., & Mahone, E. M. (2010). Factor structure of paediatric timed motor examination and its relationship with IQ. *Developmental Medicine and Child Neurology*, *52*, e188–e194. doi:10.1111/j.1469-8749.2010.03670.x

McCarthy, D. (1972). *McCarthy Scales of Children's Abilities*. New York, NY: Psychological Corporation.

Miller, D. I., & Halpern, D. F. (2014). The new science of cognitive sex differences. *Trends in Cognitive Sciences*, *18*, 37–45. doi:10.1016/j.tics.2013.10.011

Naglieri, J. A., & Bardos, A. N. (1987). *Draw-a-Person and matrix analogies test cross-cultural validity*. Paper presented at the annual meeting of the National Association of School Psychologists. New Orleans, LA.

Neale, M. C., Boker, S. M., Bergeman, C. S., & Maes, H. H. (2005). The utility of genetically informative data in the study of development. In S. M. Boker & C. S. Bergeman (Eds.), *Methodological issues in aging research* (pp. 269–327). Mahwah, NJ: Erlbaum.

nFerNelson. (1999). *Mathematics 5–14 series*. London, UK: Author.

Oliver, B., Dale, P., Saudino, K., Petrill, S., Pike, A., & Plomin, R. (2002). The validity of a parent-based assessment of cognitive abilities in three-year olds. *Early Child Development and Care*, *172*, 337–348. doi:10.1080/03004430212713

Oliver, B., Harlaar, N., Hayiou-Thomas, M. E., Kovas, Y., Walker, S., Petrill, S., . . . Plomin, R. (2004). A twin study of teacher-reported mathematics performance and low performance in 7-year-olds. *Journal of Educational Psychology*, *96*, 504–517. doi:10.1037/0022-0663.96.3.504

Oliver, B. R., & Plomin, R. (2007). 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*, *10*, 96–105.

Pangelinan, M. M., Zhang, G., VanMeter, J. W., Clark, J. E., Hatfield, B. D., & Haufler, A. J. (2011). Beyond age and gender: Relationships between cortical and subcortical brain volume and cognitive-motor abilities in school-age children. *NeuroImage*, *54*, 3093–3100. doi:10.1016/j.neuroimage.2010.11.021

Plomin, R. (2011). Commentary: Why are children in the same family so different? Non-shared environment three decades later. *International Journal of Epidemiology*, *40*, 582–592. doi:10.1093/ije/dyq144

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2013). *Behavioral genetics* (6th ed.). New York, NY: Worth.

Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin*, *131*, 592–617. doi:10.1037/0033-2909.131.4.592

Qualifications and Curriculum Authority. (2003). *QCA key stage 1: Assessment and reporting arrangements*. London, UK: Author.

Raven, J. C., Court, J. H., & Raven, J. (1996). *Manual for Raven's Progressive Matrices and Vocabulary Scales: Section 3—Standard progressive matrices*. Oxford, UK: Oxford Psychologists Press.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Standard Progressive Matrices*. Oxford, UK: Oxford Psychologists Press.

Reynolds, C. R. (1978). The McCarthy drawing tests as a group instrument. *Contemporary Educational Psychology*, *3*, 169–174. doi:10.1016/0361-476X(78)90024-3

Rijsdijk, F. V., & Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *Briefings in Bioinformatics*, *3*, 119–133.

Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, *35*, 83–92. doi:10.1016/j.intell.2006.05.004

Rutter, M. (2006). *Genes and behaviour, nature-nurture interplay explained*. Oxford, UK: Blackwell.

Saudino, K. J., Dale, P. S., Oliver, B., Petrill, S. A., Richardson, V., Rutter, M., . . . Plomin, R. (1998). The validity of parent-based assessment of cognitive abilities of 2-year-olds. *British Journal of Developmental Psychology*, *16*, 349–363.

Tosto, M., Hanscombe, K. B., Haworth, C. M. A., Davis, O. S. P., Petrill, S. A., Dale, P. S., . . . Kovacs, Y. (2014). Why do spatial abilities predict mathematical performance? *Developmental Science*, *17*, 1–9. doi:10.1111/desc.12138

Trzaskowski, M., Davis, O. S. P., DeFries, J. C., Yang, J., Visscher, P. M., & Plomin, R. (2013). DNA evidence for strong genome-wide pleiotropy of cognitive and learning abilities. *Behavior Genetics*, *43*, 267–273. doi:10.1007/s10519-013-9594-x

van Garderen, D. (2006). Spatial visualization, visual imaginary, and mathematical problem solving of students with varying abilities. *Journal of Learning Disabilities*, *39*, 496–506.

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

Voyer, D., Potsma, A., Brake, B., & Imperato-McGinley, J. (2007). Gender differences in object location memory: A meta-analysis. *Psychonomic Bulletin & Review*, *14*, 23–38.

Wang, Z., Hart, S. A., Kovas, Y., Lukowski, S., Soden, B., Thompson, L. A., . . . Petrill, S. A. (2014). Who is afraid of math? Two sources of genetic variance for mathematical anxiety. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *55*, 1056–1064. doi:10.1111/jcpp.12224

Wechsler, D. (1992). *Wechsler Intelligence Scale for Children–third ed., U.K. (WISC–III UK) Manual*. London, UK: The Psychological Corporation.

Wei, W., Yuan, H., Chen, C., & Zhou, X. (2011). Cognitive correlates of performance in advanced mathematics. *British Journal of Educational Psychology*, *82*, 157–181. doi:10.1111/j.2044-8279.2011.02049.x

Willcock, E., Imuta, K., & Hayne, H. (2011). Children's human figure drawings do not measure intellectual ability. *Journal of Experimental Child Psychology*, *110*, 444–452. doi:10.1016/j.jecp.2011.04.013

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S. D., Henders, A. K., Nyholt, D. R., . . . Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, *42*, 565–569.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Table S1. Inter-Rater Agreement Between 3 Raters for Each Item of the Drawing Maturity Scale (DMS)

Table S2. Loadings on the Two Main Components of the Drawing Maturity Scale (DMS)

Table S3. Internal Validity of the Drawing Maturity Scale

Table S4. Between-Participants ANOVAs for Gender Across the Three Drawing Measures

Table S5. Results for Four Linear Regressions with Drawing Maturity Scale (DMS) at 4 and *g* at 12 as the Predictors of Mathematics at 12

Table S6. MZ and dizygotic (DZ) Correlations and Univariate ACE Estimates (With 95% CIs) for All Variables

Table S7. Standardized Squared Paths for the Association Between *g* at Age 4, Drawing at Age 4 and Teacher-Rated Mathematics at Age 12; Goodness of Fit Indices

Table S8. Standardized Squared Paths for the Association Between g at Age 4, Drawing at Age 4 and Mathematics Web Test at Age 12

Table S9. Phenotypic (r_P), Genetic (rG), Shared (rC) and Non-Shared Environmental (rE) Correlations and Bivariate Heritability/Environmentalities for the Trivariate Association Between *g* at 4, Drawing at 4, and Web-Measured Mathematics at 12

Table S10. Standardized Squared Paths for the Association Between Drawing at Age 4, *g* at Age 12, and Teacher-Rated Mathematics at Age 12

Table S11. Phenotypic (r_P), Genetic (rG), Shared (rC), and Non-Shared Environmental (rE) Correlations and Bivariate Heritability/Environmentalities for the Trivariate Association Between, Drawing at 4, *g* at 12, and Web-Measured Mathematics at 12

AppendixS1. Drawing Maturity Scale (DMS) Validation

Appendix S2. The Drawing Maturity Scale

Appendix S3. Between-Participants ANOVAs for Gender Across the Three Drawing Measures

Appendix S4. Trivariate ACE Estimates for the Model Including *g* at 4, Drawing at 4, and Teacher-Rated Mathematics at 12

Appendix S5. Trivariate ACE Estimates for the Model Including *g* at 4, Drawing at 4, and Web-Measured Mathematics at 12

Appendix S6. Trivariate ACE Estimates for the Model Including Drawing at 4, *g* at 12, and Teacher-Rated Mathematics at 12