Published in final edited form as:

J Child Psychol Psychiatry. 2009 October ; 50(10): 1318–1325. doi:10.1111/j.1469-7610.2009.02114.x.

Generalist genes and learning disabilities: a multivariate genetic analysis of low performance in reading, mathematics, language and general cognitive ability in a sample of 8000 12-year-old twins

Claire M.A. Haworth¹, Yulia Kovas², Nicole Harlaar³, Marianna E. Hayiou-Thomas⁴, Stephen A. Petrill³, Philip S. Dale⁵, and Robert Plomin¹

¹Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, UK

²Psychology Department, Goldsmiths College, University of London, UK

³Department of Human Development and Family Science, Ohio State University, USA

⁴Department of Psychology, University of York, UK

⁵Department of Speech and Hearing Sciences, University of New Mexico, USA

Abstract

Background—Our previous investigation found that the same genes influence poor reading and mathematics performance in 10-year-olds. Here we assess whether this finding extends to language and general cognitive disabilities, as well as replicating the earlier finding for reading and mathematics in an older and larger sample.

Methods—Using a representative sample of 4000 pairs of 12-year-old twins from the UK Twins Early Development Study, we investigated the genetic and environmental overlap between internet-based batteries of language and general cognitive ability tests in addition to tests of reading and mathematics for the bottom 15% of the distribution using DeFries-Fulker extremes analysis. We compared these results to those for the entire distribution.

Results—All four traits were highly correlated at the low extreme (average group phenotypic correlation=0.58) and in the entire distribution (average phenotypic correlation=0.59). Genetic correlations for the low extreme were consistently high (average=0.67), and non-shared environmental correlations were modest (average=0.23). These results are similar to those seen across the entire distribution (0.68 and 0.23, respectively).

Conclusions—The 'Generalist Genes Hypothesis' holds for language and general cognitive disabilities, as well as reading and mathematics disabilities. Genetic correlations were high indicating a strong degree of overlap in genetic influences on these diverse traits. In contrast, non-shared environmental influences were largely specific to each trait; causing phenotypic differentiation of traits.

Keywords

learning disability; twins; behavior genetics; generalist genes

Address for Correspondence: Claire M.A. Haworth, SGDP Centre, Institute of Psychiatry, King's College London, De Crespigny Park, London SE5 8AF, UK. Tel: +44(0)2078485621 Fax: +44(0)2078480866, Claire.Haworth@iop.kcl.ac.uk.

Conflict of Interest: The authors declare no conflict of interest.

Multivariate genetic analyses consistently indicate substantial genetic overlap between learning abilities such as reading and mathematics, between cognitive abilities such as verbal and spatial, and between learning and cognitive abilities (Plomin and Kovas, 2005). This finding has been referred to as the ‘Generalist Genes Hypothesis’ in the sense that the genes that affect one ability are largely the same genes that affect other abilities, even though there are some genetic effects that are specific to each ability. The purpose of the present study is to test the Generalist Genes Hypothesis for *disabilities* (low performance) rather than for abilities (normal range of variation). The only previous multivariate genetic analyses of learning disabilities involved two studies of reading and mathematics disabilities, including a report on the current sample at 10 years, which reported genetic correlations greater than 0.50 using DF extremes analysis (Knopik, Alarcón, and DeFries, 1997; Kovas et al., 2007b). The present study includes the first investigation of the etiological overlap between language disability and other learning disabilities, and also the genetic and environmental overlap between learning disabilities and general cognitive disability.

The key statistic from multivariate genetic analysis relevant for this hypothesis is the *genetic correlation* which indexes the extent to which genetic influences on one trait also affect another trait (Neale and Maes, 2003). A high genetic correlation between two traits implies that if a genetic polymorphism is associated with one trait, there is a good chance that variation in this gene would also be associated with the other trait. Genetic correlations are consistently greater than 0.50 and often near 1.0 between learning abilities (Plomin and Kovas, 2005), between cognitive abilities (Deary, Spinath, and Bates, 2006; Plomin and Spinath, 2002), and between learning and cognitive abilities (Davis et al., 2008; Haworth, Kovas, Dale, and Plomin, 2008; Kovas, Haworth, Dale, and Plomin, 2007a). For the purposes of this investigation we define domains directly relevant to school subjects as learning abilities (reading, mathematics and language) in contrast to cognitive abilities (verbal and spatial) although we recognize that the distinction is to some extent semantic.

The Generalist Genes Hypothesis predicts that genetic correlations exceed 0.50 between these four disability domains (reading, mathematics, language and general cognitive ability). We also compare multivariate genetic results for disabilities to results for abilities. Although we have previously reported multivariate genetic results for the full range of abilities using a latent variable approach that uses test level data, with the same sample and measures (Davis et al., in press), here we report alternative analyses of composite abilities (e.g., all the reading measures combined into one variable) that are more comparable to the present analyses of disabilities.

Methods

Sample

The sampling frame was the Twins Early Development Study (TEDS), a study of twins born in England and Wales in 1994–1996 (Oliver and Plomin, 2007). The TEDS sample is reasonably representative of the general population in terms of parental education, ethnicity and employment status (Kovas et al., 2007a). Zygosity was assessed through a parent questionnaire of physical similarity, which is over 95% accurate when compared to DNA testing (Price et al., 2000). For cases where zygosity was unclear, DNA testing was conducted.

We obtained informed consent for this study from the parents of the twins. At age 12, the twins completed our internet cognitive test battery (Haworth et al., 2007). 10,875 individuals participated (completed at least one test) in our battery. The mean age of the twins was

11.56 (sd=0.69). In Table 4 we present the number of complete twin pairs by zygosity for each domain.

Measures

We have previously shown that our internet-based cognitive test battery is a reliable and valid method for collecting cognitive data on children as young as 10 years old (see Haworth et al., 2007). At 12 years we obtained measures relevant for four major cognitive domains: general cognitive ability, reading, mathematics, and language. Within each broad domain, tests load highly and to a similar degree on the latent composite (Davis et al. in press); therefore we created composite scores for each of these domains. The use of composites greatly simplifies the present analyses which are based on a series of bivariate comparisons.

General cognitive ability (g)

The twins were assessed on two verbal tests, WISC-III-PI Multiple Choice Information (General Knowledge) and Multiple Choice Vocabulary subtests (Wechsler, 1992), and two non-verbal reasoning tests, the WISC-III-UK Picture Completion (Wechsler, 1992) and Raven's Standard and Advanced Progressive Matrices (Raven, Court, and Raven, 1998). We calculated a mean composite *g* scale using these four tests.

Reading

Four measures were used: two measures of reading comprehension and a measure of reading fluency presented on the internet, and a fourth measure administered over the telephone.

Reading comprehension

The twins completed an adaptation of the reading comprehension subtest of the Peabody Individual Achievement Test (PIAT_{rc}; Markwardt, 1997). The PIAT_{rc} assesses literal comprehension of sentences. The internet-based adaptation contained the same practice items, test items and instructions as the original published test.

In addition, we assessed reading comprehension using the GOAL Formative Assessment in Literacy for Key Stage 3 (GOAL plc, 2002). The GOAL is a test of reading achievement that is linked to the literacy goals in Key Stage 3 of the UK National Curriculum. Correct answers were summed to give a total comprehension score.

Reading fluency

Reading fluency was assessed using an adaptation of the Woodcock-Johnson III Reading Fluency Test (Woodcock, McGrew, and Mather, 2001) and the Test of Word Reading Efficiency (TOWRE, form B; Torgesen, Wagner, and Rashotte, 1999). The Woodcock-Johnson is a measure of reading speed and rate that requires the ability to read and comprehend simple sentences quickly e.g. "A flower grows in the sky? - Yes/No". The online adaptation consists of 98 yes/no statements; children need to indicate yes or no as quickly as possible. There is a time limit of 3 minutes for this test. Correct answers were summed to give a total fluency score.

The TOWRE, a standardized measure of fluency and accuracy in word reading skills, includes two subtests: A graded list of 85 words, called Sight-word Efficiency (SWE), which assesses the ability to read aloud real words; and a graded list of 54 non-words, called Phonemic Decoding Efficiency (PDE), which assesses the ability to read aloud pronounceable printed non-words (Torgesen et al., 1999). The child is given 45 seconds to read as many words as possible. Twins were assessed by telephone using test stimuli that had been mailed to families in a sealed package with separate instructions that the package

should not be opened until the time of testing. We calculated a mean composite reading scale using these four tests.

Mathematics

To assess mathematics, we developed a internet-based battery based on the National Foundation for Educational Research 5–14 Mathematics Series, which is linked closely to curriculum requirements in the UK (NFER-Nelson, 1999). The items were drawn from the following three categories: Understanding Number, Non-Numerical Processes and Computation and Knowledge. The mathematics battery is described in more detail elsewhere (Kovas, Haworth, Petrill, and Plomin, 2007). We calculated a mean composite mathematics scale using these three tests.

Language

To assess receptive spoken language, standardized tests were selected that could both discriminate children with language disability and reflect individual differences across the full range of ability. Furthermore, an aspect of language that becomes increasingly important in adolescence – and which shows interesting variability at this age -- is metalinguistic ability, knowledge about language itself (Nippold, 1998). For this reason, the three measures selected for testing included one designed to assess syntax, with low metalinguistic demands (*Listening Grammar*) and two with higher metalinguistic demands that assess semantics (*Figurative Language*) and pragmatics (*Making Inferences*).

Syntax

Syntax was assessed using the *Listening Grammar* subtest of the Test of Adolescent and Adult Language (TOAL-3) (Hammill et al., 1994). This test requires the child to select two sentences that have the same meaning, out of three options. The sentences are presented orally only.

Semantics

Semantics were assessed using Level 2 of the *Figurative Language* subtest of the Test of Language Competence (Wiig et al., 1989), which assesses the interpretation of idioms and metaphors; correct understanding of such non-literal language requires rich semantic representations. The child hears a sentence orally and chooses one of four answers, presented in both written and oral form.

Pragmatics

Level 2 of the *Making Inferences* subtest of the Test of Language Competence (Wiig et al., 1989) assessed an aspect of pragmatic language, requiring participants to make permissible inferences on the basis of existing (but incomplete) causal relationships presented in short paragraphs. The child hears the paragraphs orally and chooses two of four responses, presented in both written and oral form. We calculated a mean composite language scale using these three tests.

Analyses

We performed univariate and bivariate DF extremes analyses to investigate the genetic and environmental correlations for low performance. Our goal in these analyses was to clarify the etiology of low ability, not the etiology of *individual differences* among low performers. These genetic and environmental estimates from univariate and bivariate DF extremes

analyses were compared to univariate and bivariate individual differences analyses for the entire distribution.

Extremes analysis

Individuals scoring in the bottom 15% of the distribution were classified as low performers (probands). The bottom 15% was used because this is within the estimated population rates for mathematics and reading disabilities, and it provides a balance between power and selection for low ability. (NB. Results were similar for more extreme cut-offs – details available from first author). Probandwise concordance (number of probands in concordant pairs as a ratio of the total number of probands) was calculated which indicates the risk that a co-twin of a proband also meets criteria for low performance. Greater MZ than DZ concordances suggest genetic influence, but unlike twin correlations, twin concordances cannot be used to estimate genetic and environmental parameters because they do not in themselves include information about population incidence.

Rather than dichotomizing each trait for ‘cases’ versus ‘controls’ and analyzing concordance (or polychoric correlations in the case of liability-threshold models), we used DeFries-Fulker (DF) extremes analysis (DeFries and Fulker, 1988), which incorporates quantitative trait information from the co-twins of selected probands. DF extremes analysis assesses twin similarity as the extent to which the mean standardized quantitative trait score of co-twins of the selected extreme probands is below the population mean and approaches the mean standardized score of those probands (see Plomin and Kovas, 2005). This measure of twin similarity is called a *group* twin correlation (or *transformed co-twin mean*) in DF extremes analysis because it focuses on the mean quantitative trait score of co-twins rather than individual differences. Genetic influence is implied if group twin correlations are greater for MZ than for DZ twins, that is, if the mean standardized score of the co-twins is lower for MZ pairs than for DZ pairs. Doubling the difference between MZ and DZ group twin correlations estimates the genetic contribution to the average phenotypic difference between the probands and the population. The ratio between this genetic estimate and the phenotypic difference between the probands and the population is called *group heritability*. It should be noted that group heritability does not refer to individual differences among the probands – the question is not why one proband scores slightly lower than another but rather why the probands as a group have lower scores than the rest of the population.

Although DF extremes group heritability can be estimated by doubling the difference in MZ and DZ group twin correlations (Plomin, 1991), DF extremes analysis is more properly conducted using a regression model (DeFries and Fulker, 1988). The DF extremes model fits standardized scores for MZ and DZ twins to the regression equation:

$$C(X)=B_1P(X)+B_2R+A$$

where the co-twins’ scores ($C(X)$) are predicted from the probands’ scores ($P(X)$) and the coefficient of relatedness (R), which is 1.0 for MZ (genetically identical) and 0.5 for DZ twins (who are on average 50% similar genetically), and A is the regression constant. B_1 is the partial regression of the co-twin score on the proband, an index of average MZ and DZ twin resemblance independent of B_2 . The focus of DF extremes analysis is on B_2 . B_2 is the partial regression of the co-twin score on R independent of B_1 . It is equivalent to twice the difference between the means for MZ and DZ co-twins adjusted for differences between MZ and DZ probands (i.e., scores are standardized based on proband means, so that the population mean is 0 and the proband mean is 1). In other words, B_2 is the genetic contribution to the phenotypic mean difference between the probands and the population. Group heritability is estimated by dividing B_2 by the difference between the means for

probands and the population. In DF extremes analysis, group shared environmental influences are estimated as the difference between the MZ transformed co-twin mean and group heritability. Group non-shared environmental influences explain the rest of the mean difference between the probands and the population.

In contrast to univariate DF extremes analysis which selects probands as extreme on X and compares the quantitative scores of their MZ and DZ co-twins on the same variable X, bivariate DF extremes analysis selects probands on X and compares the quantitative scores of their co-twins on another variable Y, a cross-trait twin group correlation. This can be done using the same regression procedure above, but replacing C(X) with C(Y). Bivariate group heritability is the ratio between this genetic estimate and the phenotypic difference between the probands on trait X and the population on X. Bivariate group heritability provides a standardized index of the extent to which the deficit of probands for X is due to genetic factors that also influence Y. Unlike bivariate analysis of individual differences in unselected samples, such as those mentioned above, bivariate DF extremes analysis is directional in the sense that selecting probands on X and examining quantitative scores of co-twins on Y could yield different results as compared with selecting probands on Y and examining quantitative scores of co-twins on X. A group genetic correlation, the bivariate statistic of most interest for evaluating the Generalist Genes Hypothesis for low performance, can be derived from four group parameter estimates: bivariate group heritability estimated by selecting probands for X and assessing co-twins on Y, bivariate group heritability estimated by selecting probands for Y and assessing co-twins on X, and univariate group heritability estimates for X and for Y:

$$r_{g(xy)} = \sqrt{\frac{(B_{2xy})(B_{2yx})}{(B_{2x})(B_{2y})}}$$

where B_{2xy} is the group heritability from x to y (e.g., from g to reading) and B_{2yx} is the group heritability from y to x (e.g., from reading to g), B_{2x} is the group heritability of x (e.g., univariate group heritability of g) and B_{2y} is the group heritability of y (e.g., univariate group heritability of reading) (see Knopik et al., 1997 for further details).

Individual differences analysis

Standard twin analyses were used to estimate the genetic and environmental etiology of the continuous measures for the entire distribution of individual differences. All measures were residualised for the effects of age and sex, as is standard in twin analyses (McGue and Bouchard, Jr., 1984). Intraclass twin correlations were calculated, and univariate model-fitting analyses were performed using raw data in Mx (Neale et al., 2006). Multivariate twin analyses that decompose the covariance between traits into genetic and environmental parameters were performed using a four-variable Cholesky decomposition model, and results were transformed to provide a correlated factors solution (Neale & Maes, 2003). The correlated factors solution provides estimates of the genetic and environmental overlap between the four traits, represented as genetic and environmental correlations. Means and standard deviations for these measures have been published previously (Davis et al., in press).

Results

Learning and cognitive disabilities

Probandwise concordances and group twin correlations were higher in MZ than DZ pairs for each scale, indicating that genetic influences are important (Table 1). Univariate group

heritability estimates (h^2g in Table 1) all exceed 50%, whereas the group shared environmental estimates (c^2g) are modest (0%–13%). Results from the bivariate DF extremes analyses (performed in both directions for each pair of scales) can be found in Table 2. The last column in Table 2 indicates the group phenotypic correlation. Using the first row of Table 2 as an example, the group phenotypic correlation indicates the extent to which reading scores for probands selected for low 'g' regress back to the population mean. The average of the group phenotypic correlations in Table 2 is 0.58, indicating substantial phenotypic overlap between learning disabilities. MZ cross-concordances (column 1) and MZ twin group cross-correlations (column 3) were greater than for DZ pairs (columns 2 and 4), suggesting genetic contributions to the group phenotypic correlations, which is confirmed by the bivariate DF genetic estimates (column 5, A). For the first row in Table 2, the group bivariate genetic estimate of 0.40 is 77% of the group phenotypic correlation of 0.52; 77% is the estimate for the proportion of the group phenotypic correlation that is explained by genetic influences (column 8, A/Grp). This last statistic is a proportional measure of how much the relationship between poor performance in two domains is due to common genetic influence.

Group genetic and environmental correlations were calculated from these bivariate DF extremes analyses (Table 3) using the equation described in Knopik et al. (1997). Group genetic correlations were high (0.67 on average), and particularly high between g and math (0.89), and between g and language (0.80). The lowest genetic correlation of 0.44 was found between reading and language. Shared environmental correlations were also high, with the exception of correlations with reading performance. The zero correlations with reading performance were due to the univariate group shared environment for reading being zero (Table 1). Although the remaining shared environmental correlations were all high (average=0.86), indicating overlap in the shared environmental influences on these disabilities, the univariate shared environmental contribution to all of these disabilities is minimal (average=0.09). Finally, non-shared environmental group correlations were consistently modest across all of the disabilities (average=0.23).

Analyses of individual differences for the entire sample

In order to compare these genetic and environmental correlations for learning disabilities with those for the normal distribution, we conducted multivariate genetic analyses of the entire sample. Table 4 presents univariate results for the four measures including twin intraclass correlations and model-fitting results. The individual differences heritability estimates in Table 4 (average=0.51) are consistently lower than the group heritability estimates in Table 1 (average=0.61). For reading and language, the confidence intervals for the two types of heritability do not overlap, suggesting that genetic influence is significantly greater for reading and language disabilities than for abilities.

Table 5 summarizes results from standard multivariate twin analyses for the entire sample. The phenotypic correlations are quite similar to the group phenotypic correlations in Table 2: The average phenotypic correlation in Table 5 is 0.59 and the average group phenotypic correlation in Table 2 is 0.58. Bivariate heritabilities (column 4 in Table 5) are generally lower than the group bivariate heritabilities in Table 2 (0.58 and 0.69, respectively), which might reflect the lower individual differences heritabilities as compared to group heritabilities (Table 4 and Table 1, respectively).

Finally, Table 5 presents genetic and environmental correlations for the entire sample. The genetic correlations for abilities throughout the distribution are highly similar to those for disabilities; the average genetic correlations are 0.68 for abilities and 0.67 for disabilities.

Discussion

For the first time, we show that genetic correlations are substantial not only between reading and mathematics disabilities (0.61) but also between reading and language (0.44) and between mathematics and language (0.67). It is interesting that genetic correlations between learning disabilities and general cognitive disability are even greater than genetic correlations between learning disabilities: 0.63 between reading disability and general cognitive disability, 0.89 between mathematics disability and general cognitive disability, and 0.80 between language and general cognitive disability. Finally, we show that genetic correlations for disabilities (lowest 15%) are similar to those for abilities (entire sample) using the same large sample and same measures. Between learning disabilities the average genetic correlation is 0.57; between learning abilities the average genetic correlation is 0.62. Between learning disabilities and general cognitive disability the average genetic correlation is 0.77; between learning abilities and general cognitive ability the average genetic correlation is 0.72. The similarity in results for disability and ability suggests, but does not prove, that the same generalist genes affect disabilities and abilities.

These multivariate results strongly support the Generalist Genes Hypothesis for disabilities as well as abilities. The ultimate test of the Generalist Genes Hypothesis will come when genes are found that are associated with learning and cognitive disabilities and abilities. The hypothesis predicts that most genes associated with any learning or cognitive disabilities or abilities will be associated with all of them, even though some genes will have more specific effects. The far-reaching implications of the Generalist Genes Hypothesis for molecular genetics and neuroscience have been discussed elsewhere (Kovas and Plomin, 2006; Plomin, Kovas and Haworth, 2007). For clinical child psychology and psychiatry, one obvious implication involves diagnosis: Learning disabilities are not distinct diagnostic categories from a genetic perspective. Although causes do not necessarily relate to cures, the genetic overlap between learning disabilities could have implications for treatments and prevention strategies that target this genetic overlap.

We have focused on the genetic aspect of the results in order to address the Generalist Genes Hypothesis, but these analyses also yield interesting results in relation to the environment. Although shared environmental influence is modest (0.00–0.13, Table 1), shared environmental group correlations between disabilities (Table 3) are very high except for comparisons involving reading which shows no shared environmental influence. In other words, to the extent that there are shared environmental influences, these factors operate almost entirely as generalists. In contrast, although non-shared environment accounts for more variance (0.26–0.35, Table 1), non-shared environmental correlations are low (0.16–0.29, Table 3). This pattern of results suggests that non-shared environmental influences operate as specialists that are independent for learning and cognitive abilities. Although non-shared environment includes error of measurement, a latent factor approach to normal variation using this same sample and same measures suggests that non-shared environment is largely responsible for dissociations even taking error of measurement into account (Davis et al., in press). These environmental findings for learning disabilities are similar to the results from our analyses of normal variation (Table 5): The average correlations are 0.94 for shared environment and 0.23 for non-shared environment. These findings suggest that, for disabilities as well as abilities, shared environments are almost entirely generalists and non-shared environments are largely specialists. One direction for future research is to identify the environmental factors responsible for these generalist and specialist effects and to use these factors in prevention strategies.

Limitations of these analyses include the usual limitations of the classical twin design and its extension to multivariate genetic analysis (Plomin et al., 2008). In addition there are four

potential limitations specific to our study. First, we used a moderate cut-off of 15% in order to balance selection for low performance against statistical power; however, when we conducted the same analyses using a cut-off of 5% we found similar results (details available from first author). Second, the age of our sample was 12 years; these results may not generalize to older samples. An interesting direction for future research is to explore the developmental course of generalist genes. Third, a potential limitation is that our results are based on tests administered via the internet; however previous research indicates that these internet-based tests are reliable and valid as compared to in-person testing (Haworth et al., 2007). Finally, as with any statistical procedure, bivariate DF extremes analysis has limitations; for example, analyses must be done pairwise rather than including all four variables in the same model. Nevertheless, when we applied an alternative procedure, liability-threshold model fitting, to the dichotomous data used to calculate the concordances in Table 1 and Table 2, we found similar results supportive of the Generalist Genes Hypothesis (details available from first author).

We conclude that genetic influences are largely general for learning disabilities (reading, mathematics and language) and that these generalist genes for learning disabilities are even more general in that they also affect general cognitive disability. In contrast, non-shared environmental influences are largely specialists.

Key Points

- A previous investigation reported in this journal found that most of the same genes influence poor reading and mathematics performance in 10-year-old twins.
- Here we investigate for the first time whether this finding extends to language and general cognitive ability in an older and larger sample of twins.
- Because genetic correlations were high, we conclude that there is a strong degree of overlap in genetic influences across disabilities as well as abilities, providing support for the Generalist Genes Hypothesis.
- We conclude that ostensibly distinct learning disabilities such as language, reading, and mathematics are not distinct diagnostic categories from a genetic perspective.
- The genetic overlap between learning disabilities could motivate treatment and prevention strategies, as well as molecular genetic studies, that target this genetic overlap.

Abbreviations

TEDS	Twins Early Development Study
ACE	Additive genetic (A), Shared environment (C), and non-shared environment (E)
MZ	Monozygotic
DZ	Dizygotic
DF extremes analysis	DeFries-Fulker extremes analysis

Acknowledgments

We gratefully acknowledge the ongoing contribution of the TEDS families. TEDS is supported by a program grant (G0500079) from the U.K. MRC; our work on school environments and academic achievement is also supported by grants from the US NIH (HD44454 and HD46167).

References

- Davis OSP, Haworth CMA, Plomin R. Generalist genes for cognition: etiology of learning abilities and disabilities in early adolescence. *Cognitive Neuropsychiatry*. (in press).
- 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.
- Deary IJ, Spinath FM, Bates TC. Genetics of intelligence. *European Journal of Human Genetics*. 2006; 14:690–700. [PubMed: 16721405]
- DeFries JC, Fulker DW. Multiple regression analysis of twin data: Etiology of deviant scores versus individual differences. *Acta Geneticae Medicae et Gemellologicae*. 1988; 37:205–216.
- GOAL plc. GOAL Formative Assessment in Literacy (Key Stage 3). London: Hodder & Stoughton; 2002.
- Hammill, DD.; Brown, VL.; Larsen, SC.; Wiederholt, JL. Test of Adolescent and Adult Language (TOAL-3). Austin, TX: Pro-Ed; 1994.
- Haworth CMA, Harlaar N, Kovas Y, Davis OSP, Oliver BR, Hayiou-Thomas ME, et al. Internet cognitive testing of large samples needed in genetic research. *Twin Research and Human Genetics*. 2007; 10:554–563. [PubMed: 17708696]
- Haworth CMA, Kovas Y, Dale PS, Plomin R. Science in Elementary School: Generalist Genes and School Environments. *Intelligence*. 2008; 36:694–701.
- 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, Haworth CMA, Dale PS, Plomin R. The genetic and environmental origins of learning abilities and disabilities in the early school years. *Monographs of the Society for Research in Child Development*. 2007a; 72:vii–160. [PubMed: 17995572]
- Kovas Y, Haworth CMA, Harlaar N, Petrill SA, Dale PS, Plomin R. Overlap and specificity of genetic and environmental influences on mathematics and reading disability in 10-year-old twins. *Journal of Child Psychology & Psychiatry*. 2007b; 48:914–922. [PubMed: 17714376]
- 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]
- Kovas Y, Plomin R. Generalist genes: Implications for cognitive sciences. *Trends in Cognitive Science*. 2006; 10:198–203.
- Markwardt, FC, Jr. Peabody Individual Achievement Test-Revised (Normative Update) Manual. Circle Pines: American Guidance Service; 1997.
- McGue M, Bouchard TJ Jr. Adjustment of twin data for the effects of age and sex. *Behavior Genetics*. 1984; 14:325–343. [PubMed: 6542356]
- Neale, MC.; Boker, SM.; Xie, G.; Maes, H. Mx: Statistical Modeling. 7th ed.. Department of Psychiatry; 2006. VCU Box 900126, Richmond, VA 23298
- Neale, MC.; Maes, H. Methodology for genetic studies of twins and families. Dordrecht, The Netherlands: Kluwer Academic Publishers B.V.; 2003.
- nferNelson. Mathematics 5–14 series. Windsor, UK: 1999.
- Nippold, MA. Later language development: The school-age and adolescent years. Austin, TX: Pro-Ed; 1998.
- Oliver BR, 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]

- Plomin, R. Genetic risk and psychosocial disorders: links between the normal and abnormal. In: Rutter, M.; Casaer, P., editors. *Biological risk factors for psychosocial disorders*. Cambridge: Cambridge University Press; 1991. p. 101-138.
- Plomin, R.; DeFries, J.C.; McClearn, G.E.; McGuffin, P. *Behavioral Genetics*. 5th Edition. New York: Worth; 2008.
- Plomin R, Kovas Y. Generalist genes and learning disabilities. *Psychological Bulletin*. 2005; 131:592–617. [PubMed: 16060804]
- Plomin R, Kovas Y, Haworth CMA. Generalist genes: genetic links between brain, mind and education. *Mind, Brain, and Education*. 2007; 1:11–19.
- Plomin R, Spinath FM. Genetics and general cognitive ability (g). *Trends in Cognitive Science*. 2002; 6:169–176.
- Price TS, Freeman B, Craig IW, Petrill SA, Ebersole L, Plomin R. Infant zygosity can be assigned by parental report questionnaire data. *Twin Research*. 2000; 3:129–133. [PubMed: 11035484]
- Raven, J.C.; Court, J.H.; Raven, J. *Manual for Raven's Progressive Matrices*. London: H.K. Lewis; 1998.
- Torgesen, J.K.; Wagner, R.K.; Rashotte, C.A. *Test of Word Reading Efficiency*. Austin, TX: Pro-Ed; 1999.
- Wechsler, D. *Wechsler intelligence scale for children-Third Edition UK (WISC-IIIUK) Manual*. London: The Psychological Corporation; 1992.
- Wiig, E.H.; Secord, W.; Sabers, D. *Test of Language Competence-Expanded Edition*. San Antonio, TX: The Psychological Corporation; 1989.
- Woodcock, R.W.; McGrew, K.S.; Mather, N. *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing; 2001.

MZ and DZ probandwise concordances and results of univariate DF extremes analysis using a 15% cut-off

Table 1

	Probandwise concordance		Twin group correlation		DF estimates		
	MZ	DZ	MZ	DZ	h^2_g (SE)	c^2_g (SE)	e^2_g (SE)
General Cognitive Ability	0.49	0.34	0.65	0.39	0.52 (.07)	0.13 (.05)	0.35 (.05)
Reading	0.58	0.31	0.74	0.37	0.74 (.06)	0.00 (.05) [†]	0.26 (.04)
Math	0.52	0.35	0.66	0.38	0.56 (.07)	0.10 (.05)	0.34 (.05)
Language	0.56	0.35	0.71	0.41	0.60 (.07)	0.11 (.05)	0.29 (.05)

Note. MZ=monozygotic; DZ=dizygotic; h^2_g =group heritability; c^2_g =group shared environment; e^2_g =group non-shared environment;

[†] non-significant. Preliminary analyses revealed no significant sex differences, so to increase power we combined same-sex and opposite-sex dizygotic twins. All parameter estimates (except c^2_g for reading) were significant at the 0.05 alpha level. For general cognitive ability, the low group included 1172 probands in 364 MZ and 572 DZ pairs; for reading 1514 probands in 426 MZ and 771 DZ pairs; for math 1388 probands in 380 MZ and 722 DZ pairs; and for language 1218 probands in 341 MZ and 615 DZ pairs.

Table 2

MZ and DZ probandwise concordances and results of bivariate DF extremes analysis using a 15% cut-off

	Probandwise cross-concordance		Twin group cross-correlation		Bivariate DF estimates			Proportion of phenotypic correlation explained			Group Phenotypic Correlation
	MZ	DZ	MZ	DZ	A (SE)	C (SE)	E (SE)	A/Grp	C/Grp	E/Grp	
g to reading	0.38	0.26	0.47	0.27	0.40 (.07)	0.06 (.05) [†]	0.06 (.05) [†]	0.77	0.12	0.12	0.52
Reading to g	0.38	0.27	0.45	0.27	0.38 (.08)	0.08 (.06) [†]	0.04 (.05) [†]	0.76	0.16	0.08	0.50
g to math	0.41	0.31	0.55	0.31	0.47 (.08)	0.08 (.06) [†]	0.10 (.05)	0.72	0.12	0.15	0.65
Math to g	0.45	0.30	0.54	0.30	0.49 (.08)	0.06 (.05) [†]	0.06 (.05) [†]	0.80	0.10	0.10	0.61
g to language	0.42	0.32	0.54	0.33	0.43 (.07)	0.11 (.05)	0.08 (.05)	0.69	0.18	0.13	0.62
Language to g	0.44	0.31	0.59	0.36	0.47 (.08)	0.12 (.06)	0.07 (.05) [†]	0.71	0.18	0.11	0.66
Reading to math	0.40	0.29	0.51	0.29	0.44 (.08)	0.07 (.05) [†]	0.09 (.05)	0.73	0.12	0.15	0.60
Math to reading	0.41	0.28	0.46	0.28	0.35 (.07)	0.11 (.05)	0.08 (.04)	0.65	0.20	0.15	0.54
Reading to language	0.37	0.30	0.46	0.30	0.32 (.08)	0.14 (.05)	0.07 (.05) [†]	0.60	0.26	0.13	0.53
Language to reading	0.38	0.29	0.47	0.34	0.27 (.07)	0.20 (.05)	0.07 (.05) [†]	0.50	0.37	0.13	0.54
Math to language	0.42	0.31	0.50	0.29	0.42 (.07)	0.08 (.05)	0.04 (.04) [†]	0.78	0.15	0.07	0.54
Language to math	0.40	0.31	0.52	0.34	0.36 (.08)	0.16 (.06)	0.10 (.05)	0.58	0.26	0.16	0.62

Note. MZ=monozygotic; DZ=dizygotic; A/Grp=proportion of the group phenotypic correlation attributable to genetic influences; C/Grp=proportion of attributable to shared environmental influences; E/Grp=proportion attributable to non-shared environmental influences;

[†]=non-significant. Group phenotypic correlations (Grp) are calculated by taking the mean proband score for y. The proportion of the phenotypic correlation that is accounted for by genetic influences (A/Grp) is calculated by dividing the shared genetic contribution (Column 5, A) by the group phenotypic correlation. The same calculations were applied to the environmental contributions.

Table 3

Genetic and environmental correlations from DF extremes analysis using a 15% cut-off

	Group genetic correlation	Group shared environment correlation	Group non-shared environment correlation
g and reading	0.63	0.00	0.16
g and math	0.89	0.61	0.22
g and language	0.80	0.96	0.23
Reading and math	0.61	0.00	0.29
Reading and language	0.44	0.00	0.25
Math and language	0.67	1.00*	0.20

* *Note.* One of the group shared environmental correlations exceeded one (calculated as 1.08 for math and language); this is a function of the equation for calculating the correlations, which does not restrict the correlation to fall between -1 and $+1$. Because correlations cannot exceed unity, we present this as a correlation of 1.0.

Table 4

Intraclass twin correlations and univariate ACE estimates from model-fitting analyses of individual differences using the entire sample

	rMZ	rDZ	A (95% CIs)	C (95% CIs)	E (95% CIs)
General cognitive ability	0.67 N=1464	0.43 N=2418	0.47 (0.40–0.55)	0.20 (0.14–0.26)	0.33 (0.30–0.35)
Reading	0.75 N=1852	0.44 N=3180	0.62 (0.57–0.67)	0.14 (0.09–0.18)	0.24 (0.22–0.26)
Math	0.67 N=1696	0.43 N=2896	0.49 (0.43–0.56)	0.19 (0.14–0.24)	0.32 (0.29–0.34)
Language	0.66 N=1537	0.44 N=2575	0.45 (0.38–0.52)	0.22 (0.16–0.28)	0.33 (0.31–0.35)

Note. rMZ=monozygotic twin correlation; rDZ=dizygotic twin correlation; A=additive genetic; C=shared environment; E=non-shared environment. N=number of complete pairs; CIs=confidence intervals.

Table 5

Results from multivariate model-fitting analyses of individual differences using the entire sample (with 95% confidence intervals)

	Genetic correlation	Shared environment correlation	Non-shared environment correlation	Bivariate A	Bivariate C	Bivariate E	Phenotypic correlation
g and reading	0.62 (0.57–0.69)	0.85 (0.67–1.00)	0.19 (0.15–0.24)	0.63 (0.54–0.72)	0.27 (0.18–0.34)	0.10 (0.07–0.13)	0.54 (0.51–0.57)
g and math	0.75 (0.70–0.84)	0.96 (0.82–1.00)	0.26 (0.21–0.29)	0.58 (0.49–0.66)	0.29 (0.22–0.36)	0.13 (0.11–0.15)	0.65 (0.62–0.67)
g and language	0.82 (0.74–0.90)	0.88 (0.75–1.00)	0.23 (0.18–0.27)	0.59 (0.51–0.68)	0.29 (0.21–0.36)	0.12 (0.09–0.15)	0.65 (0.62–0.68)
Reading and math	0.58 (0.53–0.63)	0.96 (0.79–1.00)	0.30 (0.26–0.34)	0.57 (0.50–0.66)	0.28 (0.21–0.35)	0.15 (0.12–0.17)	0.57 (0.54–0.60)
Reading and language	0.63 (0.58–0.68)	1.00 (0.89–1.00)	0.21 (0.16–0.24)	0.59 (0.53–0.67)	0.31 (0.23–0.38)	0.10 (0.08–0.13)	0.57 (0.54–0.60)
Math and language	0.65 (0.59–0.72)	0.98 (0.84–1.00)	0.19 (0.15–0.24)	0.54 (0.46–0.63)	0.35 (0.27–0.42)	0.11 (0.08–0.14)	0.58 (0.55–0.60)

Note. A=additive genetic; C=shared environment; E=non-shared environment; g=general cognitive ability. Bivariate heritability provides an estimate of the proportion of the phenotypic correlation that is accounted for by genetic influences. It is calculated by dividing the shared genetic contribution by the phenotypic correlation. The same calculations were applied to the environmental contributions.