

# Goldsmiths Research Online

*Goldsmiths Research Online (GRO)*  
*is the institutional research repository for*  
*Goldsmiths, University of London*

## Citation

Rodic, Maja; Rimfeld, Kaili; Gaysina, Daria and Kovas, Yulia. 2018. From Rare Mutations to Normal Variation: Genetic Association Study of Mathematical, Spatial, and General Cognitive Abilities. *Psychology in Russia: State of the Art*, 11(4), ISSN 2074-6857 [Article]

## Persistent URL

<http://research.gold.ac.uk/25854/>

## Versions

The version presented here may differ from the published, performed or presented work. Please go to the persistent GRO record above for more information.

If you believe that any material held in the repository infringes copyright law, please contact the Repository Team at Goldsmiths, University of London via the following email address: [gro@gold.ac.uk](mailto:gro@gold.ac.uk).

The item will be removed from the repository while any claim is being investigated. For more information, please contact the GRO team: [gro@gold.ac.uk](mailto:gro@gold.ac.uk)

## From Rare Mutations to Normal Variation: Genetic Association Study of Mathematical, Spatial, and General Cognitive Abilities

Maja Rodic<sup>a,c</sup>, Kaili Rimfeld<sup>b</sup>, Daria Gaysina<sup>d</sup> & Yulia V. Kovas<sup>a,c,\*</sup>

<sup>a</sup> Tomsk State University, Russia

<sup>b</sup> Kings College, London MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, London, UK

<sup>c</sup> Department of Psychology, Goldsmiths, University of London, UK

<sup>d</sup> EDGE Lab, School of Psychology, University of Sussex, Brighton, UK

\* Corresponding author. E-mail: y.kovas@gold.ac.uk

**Background.** Behavioral genetic findings suggest that complex traits, such as mathematical ability, general cognitive ability (intelligence; *g*), and spatial ability, are influenced by many common genetic variants of very small effects that operate across the ability continuum. Common genetic variants may also be responsible for cognitive deficits associated with rare genetic syndromes, in which whole genomic regions may be affected. To date, relatively few common genetic variants involved in cognitive traits have been identified, and these only explain a small proportion of variance in these traits.

**Objective.** The aim of the study was to find associations between mathematics-related traits and single-nucleotide polymorphisms (SNPs) within chromosomal regions involved in Williams and Prader-Willi disorders. Both disorders are characterized by patterns of weaknesses and strengths in cognitive abilities. Two types of analyses were performed (SNP-based and gene-based), using genotypic and phenotypic data available for 3000 participants from the UK.

**Results.** SNP-based tests indicated that none of the SNPs passed the demanding multiple testing correction level for any of the phenotypes. Gene-based analysis suggested that 2 pseudogenes (i.e., GOLGA8I and WHAMMP3) were significantly associated with intelligence, and 1 gene (i.e., TUBGCP5) was significantly associated with mathematics at 16 years of age.

**Conclusion.** The results are consistent with other findings demonstrating that cognitive traits are influenced by many common genetic variants with very small effects. The results also suggest that a small number of these variants may be located in the chromosomal regions affected in Prader-Willi and Williams syndrome regions.

**Key words.** Mathematical ability, spatial ability, intelligence, genetic variation, single-nucleotide polymorphism (SNP)

## Introduction

Twin studies consistently find moderate to high heritability of cognitive traits, such as mathematical ability (e.g., Polderman, Benyamin, De Leeuw, Sullivan, Van Bochoven, Visscher, & Posthuma, 2015; Tosto, Petrill, Halberda, Trzaskowski, et al., 2014; Kovas, Haworth, Dale, & Plomin, 2007; Oliver et al., 2004; Alarcón et al., 2000), general cognitive ability (Deary et al., 2006), and spatial ability (Bratko, 1996; Tosto et al., 2014). Furthermore, the recently developed Genome-wide Complex Trait Analysis (GCTA, also known as the GREML method), that derives heritability directly from DNA of unrelated individuals (Yang et al., 2011a), provides further evidence for genetic influences on cognitive traits, including different aspects of mathematics, general cognitive ability, verbal and non-verbal cognitive ability (Trzaskowski, Davis, DeFries, Visscher, & Plomin, 2013; Plomin et al., 2013). However, heritability obtained using GCTA is usually lower than that found in twin studies, as the GCTA heritability is based only on common single-nucleotide polymorphisms (SNPs), genotyped in a study, and estimates only additive genetic effects. In contrast, the twin method captures both additive and non-additive (e.g., gene-gene interactions) effects, of common and rare variants (Plomin et al., 2013).

According to the ‘generalist genes’ hypothesis (Plomin & Kovas, 2005) genes have ‘general’ effects on traits in the following 3 ways: (1) largely the same set of genes affects different aspects of cognitive ability; (2) largely the same set of genes contributes to a particular ability across development; and (3) risk variants are continuously distributed across an ability continuum contributing to individual differences.

These conclusions are based on research that found substantial genetic correlations ( $r_g$ ) — the extent to which the same genes affect different traits (Plomin, DeFries, Knopik, & Neiderhiser, 2016). For example,  $r_g > .60$  for diverse cognitive abilities have been found in twin studies (Calvin et al., 2012) and confirmed by the bivariate GCTA analysis (Trzaskowski et al., 2013). Using this population-based approach based on DNA alone, studies found a genetic correlation of .74 between mathematics and reading (Davis et al., 2009); .74 between general cognitive ability ( $g$ ) and mathematics; and .81 — between  $g$  and language (Trzaskowski et al., 2013). Studies have also found that stability across development of such traits as mathematics, literacy, science, and ( $g$ ) is largely explained by continuous genetic effects (Kovas et al., 2007; Haworth et al., 2010).

Much support has also been found for the genetic links between ability and disability (Plomin, Haworth, & Davis, 2009; Plomin & Kovas, 2005). For example, mathematical disability is not seen as a distinct category, but as a lower extreme of the distribution of mathematical ability. An individual’s position on such an ability continuum stems from a complex combination of effects of many common (and rare) DNA variants and environments (Kovas, Haworth, Dale, & Plomin, 2007; Butterworth & Kovas, 2013). Such genetic effects have been described as quantitative trait loci (QTLs), the individual DNA markers (locations) linked to complex traits. The QTL approach in molecular genetics has been applied in order to identify the multiple DNA loci associated with complex cognitive traits, such as mathematics or  $g$  (Plomin et al., 2013a).

The most widely used method for the identification of specific genetic variants implicated in complex traits is the Genome-Wide Association Study (GWAS). This is a hypothesis-free approach that allows for search for genetic markers (i.e. SNPs) of small effects across the whole genome (Hirschhorn & Daly, 2005). GWAS employs SNP microarrays (gene chips), which are used to genotype common genetic variants, usually SNPs that are found in more than 1% of the population. Many loci have already been identified as associated with cognitive traits, but individually they explain <1% of the variance (Plomin & Deary, 2015). To our knowledge, to date, there have been few GWA studies on mathematical ability (Docherty et al., 2010; Davis et al., 2014; Baron-Cohen et al., 2014; Zhu, Chen, Moyzis, Dong, & Lin, 2015) and general cognitive ability (e.g., Rietveld, Esko, Davies, Pers, Turley, Benyamin, Chabris, Emilsson, Johnson, Lee, and De Leeuw, 2014; Butcher et al., 2008; Davis et al., 2015) and none on spatial ability. The first GWAS on mathematics reported 10 nominally significant SNPs associated with mathematical variation in a sample of 10-year-olds (Docherty et al., 2010). Another study reported 5 SNPs that were associated with mathematical ability of 16-year-olds (Baron-Cohen et al., 2014). In both studies, the individual SNPs did not reach the genome-wide significance level and the effect sizes of individual SNPs were very small (i.e., .02% and OR = 1.63, respectively). The studies had relatively small sample sizes and were underpowered to detect SNPs of very small effects associated with the phenotype.

GWASs on general intelligence show a similar picture. One of the first GWASs on general intelligence (g) reported 6 nominally significant SNPs that individually explained between .1 and .4% of the variance in g, however, only 1 survived multiple comparison corrected p-value (Butcher et al., 2008). Other large studies of intelligence differences did not find genome-wide significant associations and found SNPs that explain only .2 to .5% of the variance (Rietveld et al. 2014; Benyamin et al., 2014; Desrivieres et al., 2014). Several recent studies on individual differences in how long a person stays in education (years of education) identified many DNA variants (Okbay, et al., 2016; Lee et al., 2018). These variants, discovered for years of education, are also associated with general cognitive ability and other related traits (e.g. Lee et al., 2018; Selzam et al., 2016; Okbay et al., 2016).

Although individual SNPs explain only a small amount of variance, the effects of these polymorphisms are often additive, which allows for aggregation of multiple variants into polygenic scores (Plomin & Deary, 2014). Such polygenic scores are typically referred in the literature as polygenic risk scores (PRS) or genome-wide polygenic scores (GPS). Several recent studies have shown that polygenic scores explain more variance than each variant individually, and with more SNPs included in the polygenic score, more variance in a trait is explained. For example, in contrast to individual SNPs that explained <1% of the variance, the set of 10 SNPs, found in Docherty et al. study (2010) accounted for 2.9% of the phenotypic variance in mathematical ability (Docherty et al., 2010). Moreover, one-third of the children who carried 10 or more risk alleles were nearly twice as likely to be in the lowest 15% of the distribution. Recently, a study employing

genome-wide polygenic scores, created using GWAS results on number of years of education, explained 9% of individual differences in exam performance at age 16 (Selzam et al., 2016).

Although polygenic scores already explain a substantial amount of variance in some traits, they also highlight the problem of ‘missing heritability’ - the gap between heritability estimates identified by quantitative genetic research and the GWAS-identified associations (Maher, 2008). Multiple reasons for ‘missing heritability’ have been proposed (Maher, 2008). For example, variation in complex traits results from many common variants of small effect, other types of common variants, and rare variants of small or large effects.

Although GWASs present a powerful tool for uncovering the genetic architecture of complex traits, they require very large samples in order to have the power to detect very small effects. One strategy to increase power for discovering new variants linked to cognitive traits is to focus on specific genomic regions that have been implicated in disorders characterized by specific patterns of cognitive deficits and relative advantages. This research approach is based on the assumption that common variants within these regions may be related to normal variation in cognitive ability. If this is true, then these can potentially be detected with a relatively small sample because the correction for multiple testing only involves known SNPs in these areas, rather than SNPs across the whole genome.

The current study adopts this approach and focuses on the chromosomal regions known to be affected in 2 genetic disorders: Prader-Willi syndrome (PWS) and Williams syndrome (WS). Both disorders are associated with disproportionately low mathematical skills, as well as intellectual impairment, with mean IQ scores for PWS ranging from 60 to 70 (Cassidy, 1997) and for WS ranging from 40 to 90 (Howlin, Davies, & Udwin, 1998).

PWS is a rare genetic disorder that is caused by the deletion of critical genes on the paternal copy of chromosomal region 15q11-13, with the genes on the maternal copy silenced (inactive). People with PWS show profound deficits in mathematical skills and short-term memory. All mathematical domains are affected, even beyond the impairment expected from the general cognitive level (Sulzbacher, Wong, McKeen, Glock, & MacDonald, 1981; Bertella et al., 2005). Interestingly, reading and visuo-spatial ability show relative strengths (Cassidy, 1997), with particularly good performance in solving jigsaw puzzles (Dykens, 2002).

WS is a genetic disorder caused by micro-genetic deletion on the long arm of chromosome 7 (7q11.23), with 25 to 28 genes missing from this region (Ewart et al., 1993). People with WS have poor mathematical ability, number skills, planning and problem solving, spatial working memory, and visuo-spatial long-term memory. In contrast, their spoken language is well developed (Ansari et al., 2003; Bellugi, Sabo, & Vaid, 1998; Mesulam, 1982).

We hypothesized that the genetic regions implicated in these syndromes might contain common DNA variants that are also associated with normal variation in mathematical ability, spatial ability, and general cognitive ability.

Two types of analyses were performed. We first ran an SNP association analysis across the selected regions, including all available SNPs. We then ran a gene-based analysis in which the effects of multiple SNPs were combined into genes (a functional unit of human genome) and tested for association, improving the power of the SNP-based method because it also takes into consideration the correlations among SNPs within a single gene. We tested whether the genes found in the 2 syndromic chromosomal regions show an association with the available phenotypes.

## **Methods**

### ***Participants***

The genotypic and phenotypic data for a sample of 3,152 individuals was drawn from the Twin Early Development Study (TEDS), a longitudinal study of more than 11,000 twins born in England and Wales in 1994, 1995, and 1996 (Oliver & Plomin, 2007). The data were obtained when the participants were 12 and 16 years of age. The TEDS sample is representative of the UK population in terms of demographic characteristics, based on the comparisons to the UK census data (Haworth, Davis & Plomin, 2013; Harlaar, Hayiou-Thomas, & Plomin, 2005). Data were excluded from the analyses of participants who had severe medical problems, specific chromosomal abnormalities, or had severe complications at birth. To decrease heterogeneity of ancestry (which may lead to the problem of population stratification in association studies), only data from families who declared themselves as white and whose first language was English were used for the analyses. The sample size varied depending on available data for each phenotype: from 1,112 for general cognitive ability at age 16 years to 2,906 for the composite mathematics score at age 12. The data from only 1 twin from each pair was used for the analyses to prevent phenotypic covariance of siblings affecting the estimates (Visscher, Yang, Goddard, et al. 2010).

Written parental consent was obtained before data collection and the project received approval from the Institute of Psychiatry ethics committee (05/Q0706/228).

### ***Measures***

Twelve phenotypes were tested for associations with SNPs in the 2 genetic regions: 6 mathematics-related measures (3 at age 12 and 3 at age 16), 2 tests of spatial ability, and 1 test of general cognitive ability at age 12; 2 tests of general cognitive ability and 1 test of spatial ability at age 16. The information about phenotypes and total sample sizes can be seen in Table 1 (for more detail, see Appendix 1).



Table 1  
*Phenotypic and genotypic data available*

Phenotypes assessed at age 12 (N)	Phenotypes assessed at age 16 (N)	Available SNPs
<b>Mathematics Test</b> — total score for «understanding numbers», «number and algebra», and «shape, space, and measures» (N = 2,270)	<b>Mathematics Composite</b> — composite score of problem verification and understanding number tasks (N = 1,144)	<b>26,478 SNPs</b> — Williams syndrome region (7q11.22)
<b>Mathematical Achievement</b> — teacher rated «using and applying mathematics», «number and algebra», «shape, space, & measures», and «handling data» (N = 1,705)	<b>Number Sense</b> — non-symbolic comparison of numerosity (N = 1,132)	<b>19,389 SNPs</b> — Prader-Willi syndrome region (15q11-13)
<b>Mathematics Composite</b> — an average score of mathematics test and mathematical achievement (N = 2,904)	<b>Number Line</b> — an estimation of numerical magnitudes (N = 1,257)	
<b>Jigsaw Puzzle</b> — spatial reasoning test (N = 2,627)	<b>Ravens Progressive Matrices</b> — test non-verbal (fluid) intelligence (N = 1125)	
<b>Spatial Composite</b> — first principal component of Hidden Shapes and Jigsaw (N = 2,627)	<b>Corsi Block</b> — of visuo-spatial short-term working memory (N = 1,111)	
<b>General Cognitive Ability 'g'</b> — General Knowledge, vocabulary, Ravens, and Picture Completion (N = 2,328)	<b>General cognitive ability 'g'</b> — (composite of the Vocabulary and Ravens scores) (N = 1111)	

*Note.* See Appendix 1 for the brief description of tests.

### **Available genotypic data**

Affymetrix GeneChip 6.0 DNA arrays were used for genotyping nearly 700,000 SNPs extracted from buccal cheek swabs of 3,655 unrelated individuals by Wellcome Trust Sanger Institute (Hinxton, UK) as part of the Wellcome Trust Case Consortium (See Trzaskowski et al., 2013 for full details). In addition, imputation was carried out using the IMPUTE version 2.3.0 software (Howie, Donnelly, & Marchini, 2009) on the genotype data after application of quality control procedures, using reference panel data from 1,000 genomes (Phase I, v3, build 37 (hg19)) (Siva, 2008). Of the imputed SNPs, only those that had an information score  $\geq 0.98$  were retained (see Krapohl et al., 2016 for full details). All available SNPs in the regions associated with the 2 genetic syndromes were used in the analyses: 26,478 SNPs in 7q11.22 area (WS); 19,389 SNPs in 15q11-13 area (PWS). The standard quality control criteria were applied for genotyping data, excluding those with minor allele frequency less than 1%, those with more than 2% of missing data, and those deviating from Hardy-Weinberg equilibrium ( $p < 10^{-3}$ ).

## Analyses

### *SNP-based analysis*

We performed an association analysis for SNPs in 2 genetic regions, using linear regression models under an additive genetic model, using Plink software (Purcell, 2007). Participants' sex and ancestry (first 8 principal components of the genotype data) were included as covariates in the regression model. Phenotypic data were normalized using the van der Waerden transformation (van der Waerden, 1952) due to a slight skew in the distribution of scores. The association analysis was conducted for all 12 phenotypes presented in Table 1: 6 phenotypes at age 12 (3 math tasks, 2 spatial tasks, and 1 general cognitive ability); and 6 at age 16 (3 mathematics-related phenotypes, 2 general cognitive ability, and 1 visuo-spatial ability). The results were corrected for multiple testing using Bonferroni correction, with the corrected p-values of  $1.8 \times 10^{-6}$  (0.05/26,478) for WS region and  $2.5 \times 10^{-6}$  (0.05/19,389) for the PWS region.

### *Gene-based analysis*

Gene-based analysis (VEGAS2; Mishra & Macgregor, 2014) was performed by combining the p-values of the SNPs within a gene (for all genes available in the 2 chosen genetic regions separately), to obtain an overall p-value for the association of the entire gene, accounting for the correlations between SNPs. The linkage disequilibrium (LD) was considered using the 1000 Genome II CEU (NCBI build 37) reference panel for each gene and the 50kb boundary. The statistical significance of this parameter was calculated using simulations.

This analysis leads to an increased power and has the potential to explain a greater proportion of variance (Hill, et al., 2014b; Liu, et al., 2010). In addition, gene-based analysis could be advantageous in that a gene (a functional unit of the human genome) is highly consistent across populations, unlike genetic variants, which differ across diverse human populations in allele frequencies, LD structure, and heterogeneity. Moreover, the problem of multiple-testing is reduced, in our case from several thousand SNPs to only 100 to 200 genes (Hill et al., 2016).

There were 110 genes available in the WS region and 194 genes in the Prader-Willi region. The alpha level was therefore set to be 0.00045 (i.e., 0.05/110) for the WS region and 0.00025 (i.e., 0.05/194) for the Prader-Willi region.

## Results

### *SNP-based results*

None of the SNP associations met the demanding significance criterion corrected for multiple testing for any of the phenotypes. According to the power calculation, using the Genetic Power Calculator (Purcell & Sham, 2003), there was 80% power to detect effect sizes between 1.8 and 1.9% in the largest sample available to us ( $N = 2906$ ), with p-value threshold corrected for the number of SNPs tested in WS and PWS regions. However, none of the SNPs had such (relatively) large effect sizes. This is in line with previous findings for quantitative traits with strongest as-



sociations accounting for less than 1% of the variance in height (Lango Allen et al., 2010), general cognitive ability (Meaburn, Harlaar, Craig, Schalkwyk, & Plomin, 2007), and mathematics (Docherty et al., 2010).

### **Gene-based results**

Gene based analyses showed that 1 gene and 2 pseudogenes from the Prader-Willi region (chromosome 15) reached the significance level ( $p=0.05/194=0.00025$ ) for association with phenotypes at age 16 (see Table 2).

Two pseudogenes, GOLGA8I and WHAMMP3, were both significantly associated with Ravens and general cognitive ability scores. The TUBGCP5 gene was significantly associated with composite math score.

The complete results of gene-based analyses can be seen in Appendix 2.

Table 2

*Gene-based analysis results for the Williams syndrome (7q11.22) and Prader-Willi areas (15q11-13)*

Chr	Gene ID	n SNPs	Start	Stop	Ravens	General cognitive ability	Composite Math
15	GOLGA8I	20	23205241	23312743	7.80E-05	4.00E-05	n/s
15	WHAMMP3	23	23137728	23258357	8.30E-05	4.70E-05	n/s
15	TUBGCP5	165	22783394	22923891	n/s	n/s	6.80E-05

*Note. Only genes that survived the correction for multiple testing were reported. Two genes were significantly associated with Ravens and 'g', and 1 was significantly associated with composite math score, all at age 16. Start and end positions do not include the  $\pm 50$  kb boundary.*

### **Discussion**

Molecular genetic studies to date have identified only a few SNPs, explaining the small amount of variance in mathematical and spatial ability. Moreover, much of the genetic variance in general cognitive ability also remains unidentified. To reach the statistical power necessary to detect very small effect sizes of individual SNPs, GWAS studies require very large samples. The present study aimed to identify genetic polymorphisms (SNPs) associated with cognitive traits by examining the variations within chromosomal regions implicated in Williams and Prader-Willi disorders. This approach was based on the hypothesis that genetic regions implicated in these syndromes contain genetic variants that are significantly associated with mathematical, spatial, and general cognitive abilities.

Firstly, we tested these genomic regions for single SNP associations. Secondly, we sought to increase the power to detect SNPs implicated in these cognitive traits by grouping SNPs into gene-based scores and considering the gene as a unit for the association, taking into consideration correlations among SNPs within a single gene.

The results from the first method confirm the absence of large effects of common genetic variants on any of the phenotypes included in this study. In line with previous research (Butcher, Davies, & Highton, 2006; Docherty et al., 2010), and similar to other quantitative traits, mathematical abilities, general cognitive ability, and spatial ability are likely to be affected by many genetic variants of small effects (QTLs). Because loci are expected to survive stringent corrections for multiple testing, it is likely that a large number of loci remain unidentified (e.g., considered as false negatives).

One explanation for such small (and currently undetectable) effects is that natural selection eliminates mutations of large negative effects on these fitness-relevant traits. This explanation may be relevant to our study, as all explored phenotypes have been associated with desirable life outcomes (e.g. Parsons & Bynner, 2005; Wai et al., 2009; Gottfredson & Deary, 2004;).

It is also possible that regions which contain rare mutations of large effects on cognitive traits do not contain genetic variants responsible for the normal range (Davis et al., 2015). For example, 1 study found that 40 genes related to non-syndromic autosomal recessive intellectual disability, causing extensive and deleterious effects on intelligence, are not associated with the whole range of intelligence (Reichenberg, Cederlöf, McMillan, Trzaskowski, Kapara, Fruchter, Ginat, Davidson, Weiser, Larsson, & Plomin, 2016). This suggests qualitative as opposed to quantitative differences caused by some genetic variants.

The results of the gene-based analyses indicated that on chromosome 15 (Prader-Willi region), 1 gene (TUBGCP5) was related to mathematics score, and 2 pseudogenes (GOLGA8I and WHAMMP3) were related to Ravens matrices (non-verbal ability) and general cognitive ability at age 16. Gene TUBGCP5 encodes gamma-tubulin complex-associated protein 5, which is essential for microtubule nucleation at the centrosome. The deletion of this gene and other genes (i.e., NIPA1, NIPA2, and CYFIP1) is well recognized in PWS and may be linked to observed cognitive and behavioral deficits (Elert-Dobkowska et al., 2014).

It is commonly found that diseases are associated with genetic variants from protein-coding regions; however, the majority of them (80%) fall outside coding regions, highlighting the importance of including both coding and non-coding regions in the search for genetic variants (Hindorff, et al., 2009).

Although most pseudo genes are mostly non-functional (Pink et al., 2011), 2 to 20% of them may be transcribed (Yano et al., 2004; Harrison et al., 2005; Zheng et al., 2005, 2007). For example, recurrent deletions of chromosome 15q13.3 are associated with intellectual disability, schizophrenia, autism, and epilepsy. Investigations into this region show 5 relatively recent structural configurations (.5 to .9 million years ago) ranging in size from 2 to 3 Mb. They are thought to be a result of human-specific expansions of segmental duplications in 2 independent inversion events, and all inversion breakpoints map near GOLGA8 core duplicons. Some evidence suggests that there is a mechanistic role of this core duplison and its palindromic architecture in promoting the evolutionary and disease-related instability of chromosome 15 (Antonacci et al., 2014).

## Conclusion

In the present study, none of the common genetic variants in PWS and WS regions were related to individual differences in cognition in the general population. However, the results from the gene-based analyses suggested that some variance in the cognitive traits may be explained by common SNPs in these regions.

Future studies may investigate genetic variation in regions implicated in other known genetic syndromes. Until whole genome sequencing becomes common practice, the approach of focusing on the genetic regions implicated in rare genetic disorders may be a beneficial method for genetic studies of complex traits. Future studies with larger sample sizes should also investigate other genetic variants, such as structural variants (e.g., copy-number variants), as well as gene-gene interactions. Future research should also test whether polygenic scores based on the associations observed in this study can explain more variance in cognitive phenotypes (Selzam et al., 2016; Domingue et al., 2015; de Zeeuw et al., 2014). Understanding genetic mechanisms underlying individual differences in cognition is an important step toward optimizing education for all learners.

## Limitations

Although the study was adequately powered to detect large effects, the study lacked power to detect variants of very small effect. Future studies with larger sample sizes are needed to address this issue, especially as the individual effect of SNPs are proving to be much smaller (<1%). For example, Visscher et al. (2010) suggest that a sample size of 100,000 will be needed to have the statistical power of 80% to detect genetic variants accounting for as little as 0.04% of the variance in a trait at a “genome-wide significance level” of  $p < 5 \times 10^{-8}$ .

Moreover, the approach used in the present study does not capture effects of the rarer variants (<1% of the population) of small effect, as well as those of other types of variants (e.g., structural). New gene chips are needed that will tag these variants. Future research may rely on next generation sequencing techniques, which will capture the whole DNA variation, regardless of rarity.

## Acknowledgments

We gratefully acknowledge the continuing contribution of the Twin Early Development Study (TEDS) families. We are also very grateful to the Wellcome Trust Case Consortium 2 team for their help with genotyping and quality control of our data. This research was supported by a grant from the Russian Science Foundation (project 14-48-00043) to Tomsk State University.

## References

- Alarcon, M., Knopik, V.S., & DeFries, J.C. (2000). Covariation of mathematics achievement and general cognitive ability in twins. *Journal of School Psychology, 38*, 63–77. [https://doi.org/10.1016/S0022-4405\(99\)00037-0](https://doi.org/10.1016/S0022-4405(99)00037-0)

- Ansari, D., Donlan, C., & Karmiloff-Smith, A. (2007). Atypical and typical development of visual estimation abilities. *Cortex: Special Issue on Selective Developmental Disorders*, 6, 758–768. [https://doi.org/10.1016/S0010-9452\(08\)70504-5](https://doi.org/10.1016/S0010-9452(08)70504-5)
- Ansari, D., Donlan, C., Thomas, M.S., Ewing, S.A., Peen, T., & Karmiloff-Smith, A. (2003). What makes counting count? Verbal and visuo-spatial contributions to typical and atypical number development. *Journal of Experimental Child Psychology*, 85, 50–62. [https://doi.org/10.1016/S0022-0965\(03\)00026-2](https://doi.org/10.1016/S0022-0965(03)00026-2)
- Baron-Cohen, S., Murphy, L., Chakrabarti, B., Craig, I., Mallya, U., Lakatosova, S., & Warriar, V. (2014). A Genome Wide Association Study of Mathematical Ability Reveals an Association at Chromosome 3q29, a Locus Associated with Autism and Learning Difficulties: A Preliminary Study. *Plos One*, 9(5), e96374. <https://doi.org/10.1371/journal.pone.0096374>
- Bellugi, U., Sabo, H., & Vaid, J. (1998). Spatial deficits in children with Williams syndrome. In J. Stiles-Davis, M. Kritchevsky, & U. Bellugi (Eds.), *Spatial cognition brain bases and development* (pp. 273–297): Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bertella, L., Girelli, I., Grugni, G., Marchi, S., Molinari, E., & Semenza, C. (2005). Mathematical skills in Prader-Willi syndrome. *Journal of Intellectual Disability Research*, 49, 159–169. <https://doi.org/10.1111/j.1365-2788.2004.00634.x>
- Bortolotto, Z.A., Clarke, V.R.J., Delany, C.M., Parry, M.C., Smolders, I., Vignes, M., & Collingridge, G.L. (1999). Kainate receptors are involved in synaptic plasticity. *Nature*, 402, 297–301. <https://doi.org/10.1038/46290>
- Bratko, D. (1996). Twin study of verbal and spatial abilities. *Personality and Individual Differences*, 21(4), 621–624. [https://doi.org/10.1016/0191-8869\(96\)00091-8](https://doi.org/10.1016/0191-8869(96)00091-8)
- Butcher, C., Davies, C., & Highton, M. (2006). *Designing Learning. From module outline to effective teaching*. London and New York: Routledge. <https://doi.org/10.4324/9780203968482>
- Butcher, L.M., Davis, O.S. P., Craig, I.W., & Plomin, R. (2008). Genome-wide quantitative trait locus association scan of general cognitive ability using pooled DNA and 500K single nucleotide polymorphism microarrays. *Genes, Brain and Behavior*, 7, 435–446. <https://doi.org/10.1111/j.1601-183X.2007.00368.x>
- Calvin, C., Deary, I., Webbink, D., Smith, P., Fernandes, C., Lee, S., ... & Visscher, P. (2012). Multivariate genetic analyses of cognition and academic achievement from two population samples of 174,000 and 166,000 school children. *Behavioural genetics*, 42(5), 699–710. <https://doi.org/10.1007/s10519-012-9549-7>
- Cassidy, S.B. (1997). Prader-Willi Syndrome. *Journal of Medical Genetics*. 34, 917–923. <https://doi.org/10.1136/jmg.34.11.917>
- Chabris, C.F., Hebert, B.M., Benjamin, D.J., Beauchamp, J.P., Cesarini, D., van der Loos, M.J.H.M., & Laibson, D. (2011). Most reported genetic associations with general intelligence are probably false positives. *Psychological Science*, 23(11), 1314–1323. <https://doi.org/10.1177/0956797611435528>
- Cooper, D.N., Chen, J.M., Ball, E.V., Howells, K., Mort, M., Phillips, A.D., ... & Stenson P.D. (2010). Genes, mutations, and human inherited disease at the dawn of the age of personalized genomics. *Human Mutation*, 31, 631–655. <https://doi.org/10.1002/humu.21260>
- Curfs, L.G., Wiegers, A.M., Sommers, J.R., Borghgraef, M., & Fryns, J. P. (1991). Strengths and weaknesses in the cognitive profile of youngsters with Prader-Willi syndrome. *Clinical Genetics*, 40, 430–434. <https://doi.org/10.1111/j.1399-0004.1991.tb03114.x>
- Davis, O.S.P., Butcher, L.M., Docherty, S.J., Meaburn, E.M., Curtis, C.J.C., Simpson, A., Plomin, R. (2010). A three-stage genome-wide association study of general cognitive ability:

- Hunting the small effects. *Behavior Genetics*, 40, 759–767. <https://doi.org/10.1007/s10519-010-9350-4>
- Davis, O.S., Haworth, C.M., & Plomin, R. (2009). Learning abilities and disabilities: Generalist genes in early adolescence. *Cognitive neuropsychiatry*, 14(4–5), 312–331. <https://doi.org/10.1080/13546800902797106>
- Deary, I.J., Spinath, F.M., & Bates, T.C. (2006). Genetics of intelligence. *European Journal of Human Genetics*, 14(6), 690–700. <https://doi.org/10.1038/sj.ejhg.5201588>
- Dehaene, S., Kerszberg, M., & Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences*, 95(24), 14529–14534. <https://doi.org/10.1073/pnas.95.24.14529>
- Docherty, S. J., Davis, O. P. S., Kovas, Y., Meaburn, E. L., Dale, S. P., Petrill, S. A., ... Plomin, R. (2010). A genome-wide association study identifies multiple loci associated with mathematics ability and disability. *Genes, Brain and Behaviour*, 9, 234–247.
- Dykens, E. M. (2002). Are jigsaw puzzles skills “spared” in persons with Prader–Willi syndrome? *Journal of Child Psychology and Psychiatry*, 43, 343–352. <https://doi.org/10.1111/1469-7610.00025>
- Dykens, E. M., & Kasari, C. (1997). Maladaptive behavior in children with Prader-Willi syndrome, Down syndrome, and nonspecific mental retardation. *American Journal on Mental Retardation*, 102, 228–237. [https://doi.org/10.1352/0895-8017\(1997\)102<0228:MBICWP>2.0.CO;2](https://doi.org/10.1352/0895-8017(1997)102<0228:MBICWP>2.0.CO;2)
- Ewart, A.K., Morris, C.A., Atkinson, D., Jin, W., Sternes, K., Spallone, P., ... & Keating, M.T. (1993). Hemizygoty at the elastin locus in a developmental disorder, Williams syndrome. *Nature Genetics*, 5(11), 11–16. <https://doi.org/10.1038/ng0993-11>
- Gottfredson, L.S., & Deary, I. J. (2004). Intelligence predicts health and longevity, but why? *Current Directions in Psychological Science*, 13(1), 1–4. <https://doi.org/10.1111/j.0963-7214.2004.01301001.x>
- Gross-Tsur, V., Manor, O., & Shalev, R.S. (1996). Developmental dyscalculia: prevalence and demographic features. *Dev Med Child Neurol*, 38, 25–33. <https://doi.org/10.1111/j.1469-8749.1996.tb15029.x>
- Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the ‘number sense’: The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental psychology*, 44, 1457–1465. <https://doi.org/10.1037/a0012682>
- Harlaar, N., Hayiou-Thomas, M. E., & Plomin, R. (2005). Reading and general cognitive ability: a multivariate analysis of 7-year-old twins. *Scientific Studies of Reading*, 9, 197–218. [https://doi.org/10.1207/s1532799xssr0903\\_2](https://doi.org/10.1207/s1532799xssr0903_2)
- Haworth, C.M.A., Wright, M.J., Luciano, M., Martin, N.G., De Geus, E.J.C., Van Beijsterveldt, C.E.M., ... & Plomin, R. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular psychiatry*, 15(11), 1112–1120. <https://doi.org/10.1038/mp.2009.55>
- Hirschhorn, J.N., & Daly, M.J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature reviews genetics*, 6(2), 95–108. <https://doi.org/10.1038/nrg1521>
- Hornef, N., Olbrich, H., Horvath, J., Zariwala, M.A., Fliegau, M., Loges, N.T., ... & Omran, H. (2006). DNAH5 mutations are a common cause of primary ciliary dyskinesia with outer dynein arm defects. *American Journal of Respiratory and Critical Care Medicine*, 174, 120–126. <https://doi.org/10.1164/rccm.200601-084OC>



- Howie, B. N., Donnelly, P., & Marchini, J. (2009). A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *Plos Genetics* 5. doi: e1000529.
- Howlin, P., Davies, M., & Udwin, O. (1998). Cognitive functioning in adults with Williams syndrome. *Journal of Child Psychology and Psychiatry*, 39, 183–189. <https://doi.org/10.1017/S0021963097001789>
- Huettner, J. E. (2003). Kainate receptors and synaptic transmission. *Progress in Neurobiology*, 70, 387–407. [https://doi.org/10.1016/S0301-0082\(03\)00122-9](https://doi.org/10.1016/S0301-0082(03)00122-9)
- Ibanez-Tallon, I., Pagenstecher, A., Fliegauf, M., Olbrich, H., Kispert, A., Ketelsen, U. P., ... & Omran, H. (2004). Dysfunction of axonemal dynein heavy chain Mdnah5 inhibits ependymal flow and reveals a novel mechanism for hydrocephalus formation. *Human molecular genetics*, 13, 2133–2141. <https://doi.org/10.1093/hmg/ddh219>
- Kovas, Y., Haworth, C.M.A., Dale, P.S., Petrill, S.A., & Plomin, R. (2007b). Mathematical ability of 10-year-old boys and girls: genetic and environmental etiology of typical and low performance. *Journal of Learning Disabilities*, 40, 554–567.
- Kovas, Y., Haworth, C.M.A., Dale, P.S., & Plomin, R. (2007). The genetic and environmental origins of learning abilities and disabilities in early years. *Monographs of the Society for Research in Child Development*, 72, 1–158. <https://doi.org/10.1177/00222194070400060601>
- Kovas, Y., Haworth, C.M.A., Harlaar, N., Petrill, S.A., Dale, P.S., & Plomin, R. (2007). Overlap and specificity of genetic and environmental influences on mathematics and reading disability in 10-year-old twins. *Journal of Child Psychol Psychiatry*. 48, 914–922. <https://doi.org/10.1111/j.1469-7610.2007.01748..x>
- Kovas, Y., Petrill, S.A., & Plomin, R. (2007c). The origins of diverse domains of mathematics: generalist genes but specialist environments. *Journal of Educational Psychology*, 99, 128–139. <https://doi.org/10.1037/0022-0663.99.1.128>
- Lango Allen, H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., ... & Ollander, C. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, 467, 832–838. <https://doi.org/10.1038/nature09410>
- Law, M., Stewart, D., Pollock, N., & (1998). Critical review form – Quantitative studies: McMaster University: Occupational Therapy Evidence-Based Practice Research Group.
- Lee, J.J., Wedow, R., Okbay, A., Kong, E., Maghziyan, O., et al. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, 50, 1112–1121. <https://doi.org/10.1038/s41588-018-0147-3>
- Li, M. X., Yeung, J. M., Cherny, S. S., & Sham, P. (2012). Evaluating the effective number of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Human genetics*, 13(5), 747–756. <https://doi.org/10.1007/s00439-011-1118-2>
- Maher, B. (2008). Personal genomes: The case of the missing heritability. *Nature*, 456, 18–21. <https://doi.org/10.1038/456018a>
- Marchini, J., Howie, B. N., Myers, S., McVean, G., & Donnelly, P. (2007). A new multipoint method for genome-wide association studies via imputation of genotypes. *Nature Genetics*, 38, 906–913. <https://doi.org/10.1038/ng2088>
- Meaburn, E.L., Harlaar, N., Craig, I.W., Schalkwyk, L.C., & Plomin, R. (2007). Quantitative trait locus association scan of early reading disability and ability using pooled DNA and 100K SNP microarrays in a sample of 5760 children. *Molecular Psychiatry*, 13, 729–740. <https://doi.org/10.1038/sj.mp.4002063>

- Mesulam, M. M. (1982). Slowly progressive aphasia without generalized dementia. *Annals of Neurology*, *11*, 592–598. <https://doi.org/10.1002/ana.410110607>
- Miyajima, F., Ollier, W., Mayes, A., Jackson, A., Thacker, N., Rabbitt, P., ... & Payton, A. (2008a). Brain-derived neurotrophic factor polymorphism Val66Met influences cognitive abilities in the elderly. *Genes, brain and behaviour*, *7*, 411–417. <https://doi.org/10.1111/j.1601-183X.2007.00363.x>
- Miyajima, F., Quinn, J.P., Horan, M., Pickles, A., Ollier, W.E., Pendleton, N., & Payton, A. (2008b). Additive effect of BDNF and REST polymorphisms is associated with improved general cognitive ability. *Genes, brain and behaviour*, *7*, 714–719. <https://doi.org/10.1111/j.1601-183X.2008.00409.x>
- Morris, C.A., Dempsey, S.A., Leonard, C.O., Dilts, C., & Blackburn, B.L. (1988). Natural history of Williams syndrome: Physical characteristics. *Journal of Paediatric Medicine*, *113*, 318–326. [https://doi.org/10.1016/S0022-3476\(88\)80272-5](https://doi.org/10.1016/S0022-3476(88)80272-5)
- Murphy, M. M., & Mazzocco, M. M. M. (2008). Rote numeric skills may mask underlying mathematical disabilities in girls with fragile X syndrome. *Developmental Neuropsychology*, *33*, 345–336. <https://doi.org/10.1080/87565640801982429>
- NFER. (2002). NferNelson Publishing Company Ltd.
- nferNelson. (1994). *Maths 5-14series*. London: nferNelson Publishing Company Ltd.
- O’Hearn, K., & Luna, B. (2009). Mathematical skills in Williams syndrome: insight into the importance of underlying representations. *Developmental disabilities research reviews*, *15*(1), 11–20. <https://doi.org/10.1002/ddrr.47>
- Okbay, A., J. P. Beauchamp, M. A. Fontana, J. J. Lee, T. H. Pers, C. A. Rietveld, P. Turley, et al. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*, *533* (7604): 539–542. <https://doi.org/10.1038/nature17671>
- Oliver, B., Harlaar, N., Hayiou-Thomas, M.E., Kovas, Y., Walker, S.O., Petrill, S.A., ... & 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. <https://doi.org/10.1037/0022-0663.96.3.504>
- Oliver, B., & Plomin, R. (2007). Twins’ Early Development Study (TEDS): A multivariate, longitudinal genetic investigation of language, cognition and behaviour problems from childhood through adolescence. *Twin Research and Human Genetics*, *10*, 96–105. <https://doi.org/10.1375/twin.10.1.96>
- Opfer, J.E., & Siegler, R.S. (2007). Representational change and children’s numerical estimation. *Cognitive Psychology*, *55*, 169–195. <https://doi.org/10.1016/j.cogpsych.2006.09.002>
- Paglin, M., & Rufolo, A. (1990). Heterogeneous human capital occupational choice and male-female earning differences. *Journal of labor economics*, *8*, 123–144. <https://doi.org/10.1086/298239>
- Pagulayan, F.K., Busch, R.M., Medina, K.L., Bartok, J.A., & Krikorian, R. (2006). Developmental Normative Data for the Corsi Block-Tapping Task. *Journal of Clinical and Experimental Neuropsychology*, *28*(6), 1043–1052. <https://doi.org/10.1080/13803390500350977>
- Parsons, S., & Bynner, J. (2005). Does numeracy matter more? London: National Research and Development Centre for adult literacy and numeracy.
- Paterson, S.J., Girelli, L., Butterworth, B., & Karmiloff-Smith, A. (2006). Are numerical impairments syndrome specific? Evidence from Williams syndrome and Down’s syndrome. *Journal of Child Psychology and Psychiatry*, *47*, 190–204. <https://doi.org/10.1111/j.1469-7610.2005.01460.x>

- Plomin, R. (2013). Child Development and Molecular Genetics: 14 Years Later. *Child development*, 84, 104–120. <https://doi.org/10.1111/j.1467-8624.2012.01757.x>
- Plomin, R., DeFries, J.C., Knopik, V.S., & Neiderhiser, J.M. (2016). *Behavioral genetics*. New York: Worth Publishers.
- Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin*, 131(4), 592–617. <https://doi.org/10.1037/0033-2909.131.4.592>
- Polderman, T.J., Benyamin, B., De Leeuw, C.A., Sullivan, P.F., Van Bochoven, A., Visscher, P.M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature genetics*, 47(7), 702–709. <https://doi.org/10.1038/ng.3285>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., & Ferreira, M.A.R. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 8, 81, 559–575.
- Purcell, S. C., S.S. Sham, P.C. (2003). Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics*, 19(1), 149–150. <https://doi.org/10.1086/519795>
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scales*. Oxford: Oxford Psychologists Press.
- Rietveld, C.A., Medland, S.E., Derringer, J., Yang, J., Esko, T., Martin, N.W., ... & Koellinger, P. D. (2013). GWAS of 126, 559 Individuals Identifies Genetic Variants Associated with Educational Attainment. *Science*, 340, 1467–1471.
- Sahoo, T., del Gaudio, D., German, J.R., Shinawi, M., Peters, S.U., Person, R.E., ... & Beaudet, A. L. (2008). Prader–Willi phenotype caused by paternal deficiency for the HBII-85 C/Dbox small nucleolar RNA cluster. *Nature Genetics*, 40, 719–721. <https://doi.org/10.1038/ng.158>
- Semenza, C. (2008). Number Processing. In B. Stemmer & H. Whitaker (Eds.), *Handbook of the neuroscience of language* (pp. 219–227). Oxford: Elsevier. <https://doi.org/10.1016/B978-0-08-045352-1.00021-5>
- Sulzbacher, S., Wong, B., McKeen, J., Glock, J., & MacDonald, B. (1981). Long term therapeutic effects of a three month intensive growth group. *The Journal of Clinical Psychiatry*, 42(4), 148–153.
- Thompson, L.A., Detterman, D.K., & Plomin, R. (1991). Associations between cognitive abilities and scholastic achievement: Genetic overlap but environmental differences. *Psychological Science*, 158–165. <https://doi.org/10.1111/j.1467-9280.1991.tb00124.x>
- Tosto, M.G., Petrill, S.A., Halberda, J., Trzaskowski, M., Tikhomirova, T.N., Bogdanova, O.Y., Ly, R., Wilmer, J.B., Naiman, D.Q., Germine, L., Plomin, R., & Kovas, Y. (2014). Why do we differ in number sense? Evidence from a genetically sensitive investigation. *Intelligence*, 43(100), 35–46. <https://doi.org/>
- Trzaskowski, M., Davis, O.S.P., DeFries, J.P., Visscher, P.M., & Plomin, R. (2013). DNA Evidence for strong genome-wide pleiotropy of cognitive and learning abilities. *Behavior genetics*, 43(4), 267–273. <https://doi.org/10.1007/s10519-013-9594-x>
- van der Waerden, B.L. (1952). Order tests for the two-sample problem and their power. *Indagationes Mathematicae*, 453–458. [https://doi.org/10.1016/S1385-7258\(52\)50063-5](https://doi.org/10.1016/S1385-7258(52)50063-5)
- Van Herwegen, J., Ansari, D., Xu, F., & Karmiloff-Smith, A. (2008). Small and larger number processing in infants and toddlers with Williams syndrome. *Developmental Science*, 11, 637–643. <https://doi.org/10.1111/j.1467-7687.2008.00711.x>
- Varela, M.C., Kok, F., Setian, N., Kim, S.A., & Koiffmann, C.P. (2005). Impact of molecular mechanisms, including deletion size on Prader-Willi syndrome phenotype: study of 75 patients. *Clinical Genetics*, 67, 47–52. <https://doi.org/10.1111/j.1399-0004.2005.00377.x>

- Visscher, P.M., Yang, J., & Goddard, M.E. (2010). A commentary on “Common SNPs explain a large proportion of the heritability for human height” by Yang et al. (2010). *Twin research and Human Genetics*, 13, 517–524. <https://doi.org/10.1375/twin.13.6.517>
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101(4), 817–835. <https://doi.org/10.1037/a0016127>
- Whittington, J.E., Holland, A.J., Webb, T., Butler, J., Clarke, D., & Boer, H. (2001). Population prevalence and estimated birth incidence and mortality rate for people with Prader—Willi syndrome in one UK Health Region. *Journal of Medical Genetics*, 38, 792–798. <https://doi.org/10.1136/jmg.38.11.792>
- Yang, J., Manolo, T.A., Pasquale, L.R., Boerwinkle, E., Caporaso, N., Cunningham, J.M., ... & Visscher, P.M. (2011a). Genome partitioning of genetic variation for complex traits using common SNPs. *Nature Genetics*, 43(6), 519–525. <https://doi.org/10.1038/ng.823>

*Original manuscript received August 17, 2018*  
*Revised manuscript accepted October 15, 2018*  
*First published online December 30, 2018*

## Appendix 1

### *Description of phenotypic measures*

#### *Age 12 phenotypes*

*Mathematics Test* — administered on the web, based on nferNelson (1994) Maths 5-14 Series (for details of the test see Haworth, Harlaar, Kovas, Davis, Oliver, Hayiou-Thomas, Busfield, McMillan, Dale, and Plomin, 2007).

*Mathematical Achievement* — rated by teachers, reflected children's school performance according to the UK National Curriculum criteria (for details of the test see Haworth, Harlaar, Kovas, Davis, Oliver, Hayiou-Thomas, Busfield, McMillan, Dale, and Plomin, 2007).

*Mathematics Composite* — the 3 components of the web test and the 4 components of the teacher-rated achievement were previously found to be highly correlated (e.g. Kovas, Haworth, Petrill & Plomin, 2007)

*JigsawPuzzle* — administered on the web, taken from the NferNelson Spatial Reasoning series of books (NFER, 2002, Tosto, Hanscombe, Haworth, Davis, Petrill, Dale, Malykh, Plomin & Kovas, 2014).

*Hidden Shapes* — administered on the web, taken from the NferNelson Spatial Reasoning series of books (NFER, 2002, Tosto, et al., 2014).

*SpatialComposite* — based on the Hidden Shapes and Jigsaw Puzzle items.

*General Cognitive Ability 'g'* — based on the scores of 4 tasks administered on the web: General Knowledge (Kaplan, Fein, Kramer, Delis, & Morris, 1999), Vocabulary (Kaplan, et al., 1999), Ravens (Raven, Court, & Raven, 1996), and Picture Completion (Wechsler, 1992) tests. For more information on the tests see Haworth, et al., 2007).

#### *Age 16 phenotypes*

*Mathematics Composite* — created from Problem Verification and Understanding Number Tasks (Murphy & Mazzocco 2008; nferNelson, 1994).

*Number Sense* — administered on the web, assessed the ability to discriminate numerosities (Halberda & Feigenson, 2008).

*Number Line* — administered on the web, assessed estimation of numerical magnitudes and was programmed and implemented online from a description obtained from (Opfer & Siegler, 2007).

*Ravens Progressive Matrices test* — administered on the web, a computerized test of non-verbal (fluid) intelligence, adapted from (Raven, Raven, & Court, 1998).

*General Cognitive Ability 'g'* — administered on the web, a composite of the Vocabulary (Kaplan, et al. 1999) and Ravens scores (Raven, Court, & Raven, 1996).

*Corsi Block* — administered on the web, a test of visuo-spatial short-term working memory (Pagulayan, Busch, Medina, Bartok, & Krikorian, 2006).



**Appendix 2**

*Gene-based analysis results for the William syndrome (7q11.22) and Prader-Willi areas (15q11-13). Nominally significant genes.*

Phenotype	Chr	Gene ID	nSNPs	Start	Stop	p-value
Mathematics Test (age 12)	Chr7	POMZP3	48	76189302	76306620	0.001
	Chr7	LOC100133091	92	76128657	76307299	0.004
	Chr7	POM121C	35	74996059	75165565	0.005
	Chr7	SPDYE5	46	75074298	75183628	0.006
	Chr7	PMS2P3	44	75087068	75207453	0.006
	Chr7	WBSCR28	109	73225488	73330223	0.012
	Chr7	CLDN4	116	73195192	73297023	0.015
	Chr7	WBSCR27	111	73198920	73306855	0.015
	Chr7	UPK3B	44	76089739	76207199	0.034
	Chr7	PMS2P9	10	76618796	76732355	0.045
	Chr7	NSUN5P1	18	74989604	75096071	0.045
Mathematics Composite (age 12)	Chr7	POMZP3	48	76189302	76306620	0.001
	Chr7	POM121C	35	74996059	75165565	0.003
	Chr7	LOC100133091	92	76128657	76307299	0.010
	Chr7	SPDYE5	46	75074298	75183628	0.010
	Chr7	WBSCR28	109	73225488	73330223	0.012
	Chr7	NSUN5P1	18	74989604	75096071	0.013
	Chr7	PMS2P3	44	75087068	75207453	0.014
	Chr7	WBSCR27	111	73198920	73306855	0.014
	Chr7	CLDN4	116	73195192	73297023	0.014
	Chr7	TRIM74	17	74974902	75084888	0.017
	Chr7	TRIM73	17	74974902	75084896	0.017
	Chr7	TPST1	409	65620258	65875438	0.035

Mathematical Achievement (age 12)	Chr7	TYW1	577	66411791	66754507	0.024
	Chr7	AUTS2	948	69013904	70308054	0.033
	Chr7	MIR4650-2_1	405	66529308	66629384	0.033
	Chr7	MIR4650-1_1	405	66529308	66629384	0.034
	Chr7	CALN1	570	71194475	71927360	0.035
Jigsaw Puzzle (age 12)	Chr7	LINC00174	210	65791030	65915395	0.020
	Chr7	CALN1	570	71194475	71927360	0.020
	Chr7	RABGEF1_2	427	66155642	66326448	0.022
	Chr7	LOC493754	424	65945291	66107394	0.023
	Chr7	RABGEF1_1	595	66097077	66326448	0.024
	Chr7	LOC100996437	332	66069505	66184589	0.029
	Chr7	SRCRB4D	117	75968645	76089012	0.031
	Chr7	SBDS	145	66402689	66510588	0.034
	Chr7	POMZP3	48	76189302	76306620	0.036
	Chr7	GTF2IRD1P1	282	66224979	66359813	0.036
	Chr7	TPST1	409	65620258	65875438	0.038
	Chr7	KCTD7	332	66043867	66158216	0.039
	Chr7	LOC100507468	147	69011123	69112481	0.042
	Chr7	GNAI1	267	79714139	79898725	0.044
	Spatial Composite (age 12)	Chr7	LOC100507468	147	69011123	69112481
Chr7		LINC00174	210	65791030	65915395	0.028
Chr7		LOC493754	424	65945291	66107394	0.034
g (age 12)	Chr7	POMZP3	48	76189302	76306620	0.017
	Chr7	WBSCR28	109	73225488	73330223	0.026
	Chr7	SPDYE5	46	75074298	75183628	0.033
	Chr7	POM121C	35	74996059	75165565	0.036
	Chr7	CLDN4	116	73195192	73297023	0.037
	Chr7	WBSCR27	111	73198920	73306855	0.039
	Chr7	PMS2P3	44	75087068	75207453	0.046

Number Line (age 16)	Chr7	TRIM50	76	72676531	72792085	0.008
	Chr7	NSUN5	47	72666512	72772864	0.008
	Chr7	FKBP6	85	72692154	72822646	0.008
	Chr7	MIR3914-1	311	70722657	70822756	0.012
	Chr7	MIR3914-2	311	70722659	70822754	0.013
Corsi Block (age 16)	Chr7	MDH2	211	75627336	75746827	0.003
	Chr7	STYXL1	245	75575654	75727321	0.000
	Chr7	SRRM3	222	75781210	75966609	0.002
	Chr7	PMS2P9	10	76618796	76732355	0.023
	Chr7	TMEM120A	137	75566154	75673992	0.000
	Chr7	MIR3914-2	311	70722659	70822754	0.000
	Chr7	MIR3914-1	311	70722657	70822756	0.000
	Chr7	SNORA14A	71	75523100	75623234	0.000
Ravens (age 16)	Chr7	LOC101929736	344	66750962	66855012	0.002
	Chr7	MAGI2-AS3	287	79032272	79150524	0.009
	Chr7	STAG3L4	353	66717624	66836513	0.011
	Chr7	FKBP6	85	72692154	72822646	0.035
	Chr7	TRIM50	76	72676531	72792085	0.037
	Chr7	PMS2P4	354	66691117	66817429	0.039
	Chr7	NSUN5	47	72666512	72772864	0.044
	Chr7	LOC101927243	227	76994514	77104760	0.047
g (age 16)	Chr7	LOC101929736	344	66750962	66855012	0.010
	Chr7	MAGI2-AS3	287	79032272	79150524	0.017
	Chr7	LOC101927269	101	79950832	80053755	0.027
	Chr7	STAG3L4	353	66717624	66836513	0.045
	Chr7	PMS2P9	10	76618796	76732355	0.053
Number sense (age 16)	Chr7					n.s.

Mathematics Composite (age 16)	Chr7	MAGI2-AS3	287	79032272	79150524	0.001
	Chr7	PMS2P9	10	76618796	76732355	0.003
	Chr7	MIR3914-1	311	70722657	70822756	0.003
	Chr7	MIR3914-2	311	70722659	70822754	0.003
Mathematics Test (age 12)	Chr15	OR4N3P	3	22363461	22464395	0.049
Mathematics Composite (age 12)	Chr15					n.s.
Mathematical Achievement (age 12)	Chr15	CYFIP1	131	22892648	23003603	0.033
Jigsaw Puzzle (age 12)	Chr15	GABRA5	132	27061865	27244357	0.015
Spatial composite (age12)	Chr15	TJP1	256	29942356	30164706	0.016
	Chr15	GOLGA8I	20	23205241	23312743	0.025
	Chr15	WHAMMP3	23	23137728	23258357	0.026
	Chr15	GABRA5	132	27061865	27244357	0.027
	Chr15	ATP10A	741	25873859	26158349	0.030
g (age 12)	Chr15	NIPA1	186	22993278	23136843	0.015
	Chr15	NIPA2	222	22954683	23084427	0.018
	Chr15	ATP10A	741	25873859	26158349	0.021
	Chr15	CYFIP1	349	22842648	23053603	0.041
Number Line (age 16)	Chr15	TUBGCP5	165	22783394	22923891	0.014
	Chr15	SNRPN	443	25018793	25273729	0.034
Corsi Block (age 16)	Chr15	GOLGA8N	11	32835656	32949511	0.007
	Chr15	ARHGAP11A	63	32857344	32982150	0.019
	Chr15	OTUD7A	556	31725328	31997542	0.047
Ravens (age 16)	Chr15	GOLGA8I	20	23205241	23312743	7.80E-05
	Chr15	WHAMMP3	23	23137728	23258357	8.30E-05
	Chr15	CHRNA7	311	32272685	32512384	0.019
	Chr15	ATP10A	741	25873859	26158349	0.252

g (age 16)	Chr15	GOLGA8I	20	23205241	23312743	4.00E-05
	Chr15	WHAMMP3	23	23137728	23258357	4.70E-05
	Chr15	CHRNA7	311	32272685	32512384	0.018
	Chr15	MIR211	209	31307234	31407344	0.024
Number Sense (age 16)	Chr15					n.s.
Mathematics Composite (age 16)	Chr15	TUBGCP5	165	22783394	22923891	6.80E-05
	Chr15	CYFIP1	349	22842648	23053603	0.001
	Chr15	GOLGA8I	20	23205241	23312743	0.021
	Chr15	WHAMMP3	23	23137728	23258357	0.022
	Chr15	FMN1_1	710	33007744	33536934	0.043
	Chr15	FMN1_2	359	33391984	33536934	0.049