

## Evidence of Common Genetic Overlap Between Schizophrenia and Cognition

Leon Hubbard<sup>1,6</sup>, Katherine E. Tansey<sup>1,6</sup>, Dheeraj Rai<sup>2</sup>, Peter Jones<sup>3</sup>, Stephan Ripke<sup>4</sup>, Kimberly D. Chambert<sup>4</sup>, Jennifer L. Moran<sup>4</sup>, Steven A. McCarroll<sup>4,5</sup>, David E. J. Linden<sup>1</sup>, Michael J. Owen<sup>1</sup>, Michael C. O'Donovan<sup>1</sup>, James T. R. Walters<sup>\*1,2</sup>, and Stanley Zammit<sup>1,2</sup>

<sup>1</sup>MRC Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, UK; <sup>2</sup>School of Social and Community Medicine, University of Bristol, Bristol, UK; <sup>3</sup>Department of Psychiatry, Cambridge University, Cambridge, UK; <sup>4</sup>Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA; <sup>5</sup>Department of Genetics, Harvard Medical School, Boston, MA

\*Both authors contributed equally to this work and are joint first-authors.

To whom correspondence should be addressed; MRC Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, Cardiff University School of Medicine, Hadyn Ellis Building, Maindy Road, Cardiff CF24 4HQ, UK; tel: +44-029-20-688-434, fax: +44-029-20-687-068, e-mail: [WaltersJT@cf.ac.uk](mailto:WaltersJT@cf.ac.uk)

**Cognitive impairment is a core feature of schizophrenia but there is limited understanding of the genetic relationship between cognition in the general population and schizophrenia. We examine how common variants associated with schizophrenia *en masse* contribute to childhood cognitive ability in a population-based sample, and the extent to which common genetic variants associated with childhood cognition explain variation in schizophrenia. Schizophrenia polygenic risk scores were derived from the Psychiatric Genomics Consortium ( $n = 69\,516$ ) and tested for association with IQ, attention, processing speed, working memory, problem solving, and social cognition in over 5000 children aged 8 from the Avon Longitudinal Study of Parents and Children birth cohort. Polygenic scores for these cognitive domains were tested for association with schizophrenia in a large UK schizophrenia sample ( $n = 11\,853$ ). Bivariate genome-wide complex trait analysis (GCTA) estimated the amount of shared genetic factors between schizophrenia and cognitive domains. Schizophrenia polygenic risk score was associated with lower performance IQ ( $P = .001$ ) and lower full IQ ( $P = .013$ ). Polygenic score for performance IQ was associated with increased risk for schizophrenia ( $P = 3.56\text{E-}04$ ). Bivariate GCTA revealed moderate genetic correlation between schizophrenia and both performance IQ ( $r_G = -.379$ ,  $P = 6.62\text{E-}05$ ) and full IQ ( $r_G = -.202$ ,  $P = 5.00\text{E-}03$ ), with approximately 14% of the genetic component of schizophrenia shared with that for performance IQ. Our results support the presence of shared common genetic factors between schizophrenia and childhood cognitive ability. We observe a genetic relationship between schizophrenia and performance IQ but not verbal IQ or other cognitive variables, which may have implications for studies utilizing cognitive endophenotypes for psychosis.**

**Key words:** schizophrenia/cognition/performance IQ/polygenic scoring/bivariate heritability

### Introduction

Schizophrenia is a severe psychiatric disorder characterized by positive symptoms such as delusions and hallucinations, and negative symptoms such as avolition and apathy. Deficits in cognitive ability are common and are strong predictors of functional outcome,<sup>1</sup> but are largely refractory to current treatments. Cognitive impairments often manifest premorbidly during childhood and adolescence,<sup>2</sup> with the largest effects observed for IQ.<sup>3,4</sup> These deficits persist into adulthood, typically worsen with illness onset but thereafter remain relatively static.<sup>5</sup> The Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS)<sup>6</sup> initiative identified 7 separable domains showing deficits in schizophrenia: attention, processing speed, reasoning/problem solving, social cognition, verbal learning, visual learning and working memory.

There is long-standing debate about the degree of genetic overlap between cognition in the general population and schizophrenia. The proportion of the phenotypic correlation between cognition and schizophrenia that is due to shared genetic effects was found to be substantial in a number of twin studies,<sup>7-9</sup> although estimates of the genetic variance for schizophrenia that was shared with cognition varied substantially. In the only twin study that was population-based rather than ascertained by case status, and which therefore used a measure of cognition that was ascertained premorbidly, only 7% of the genetic variance for psychosis was shared with cognition,<sup>9</sup>

compared to 21%–56% in the other studies.<sup>7,8</sup> Therefore the extent to which cognitive ability and schizophrenia share genetic aetiology remains uncertain. With recent advances in genetic technology and analytic approaches, it is now possible to investigate the underlying genetic architecture of complex traits in large cohorts of unrelated individuals.

One such approach for doing this is to use polygenic score analysis in which a score representing the additive effect of large numbers of independent (ie, in approximate linkage equilibrium) single nucleotide polymorphisms (SNPs) that are associated with 1 trait also contribute to the same or a different trait in an independent sample. Three previous studies have investigated the relationship between cognition in population samples and schizophrenia using polygenic score analysis. Increased polygenic risk of schizophrenia was weakly associated with lower general cognitive ability,<sup>10,11</sup> and with a decline in cognitive ability across the lifespan.<sup>10</sup> However these results have not been consistently replicated<sup>12</sup> and polygenic risk scores for schizophrenia explained only a small proportion of variance in cognition ( $R^2 \approx 1\%$ ). A reverse approach showed polygenic scores for general cognition ("g") were associated with schizophrenia, but explained a small proportion of variance ( $R^2 < 0.5\%$ ).<sup>11</sup> These studies all focused predominantly on "g" derived from different cognitive measures across multiple participating centers. To our knowledge, no studies to date have examined the genetic overlap between schizophrenia and a range of cognitive domains. This study aims to quantify the polygenic overlap between schizophrenia and individual cognitive domains and thus inform the debate about the strength of their genetic relatedness.

A second approach for investigating the underlying genetic architecture of complex traits is to use Genome-wide Complex Trait Analysis (GCTA) to estimate the proportion of phenotypic variance explained by common SNPs across the genome.<sup>13</sup> It does this by first determining the degree of genetic relatedness between all individuals and then estimating the proportion of phenotypic variance explained by the genetic relatedness matrix using mixed-model analysis.<sup>13</sup> The bivariate extension of GCTA allows quantification of the shared common genetic factors between 2 traits by determining the genetic variance of each trait as well as the genetic covariance, and then calculating the genetic correlation.<sup>14</sup> This method has been applied previously to show significant shared common genetic aetiology between different psychiatric disorders.<sup>15</sup> No previous studies have used bivariate GCTA to examine the amount of shared common genetic factors between cognition and schizophrenia.

In this study we use data from the largest schizophrenia genome-wide association study (GWAS) undertaken to date,<sup>16</sup> and birth cohort data on a broad range of cognitive phenotypes measured during childhood to undertake a detailed investigation of the genetic relationship

between schizophrenia and cognition using both polygenic scoring and bivariate GCTA. More specifically, we examine (1) whether common variants associated with schizophrenia *en masse* contribute to cognition in the general population at age 8, (2) the extent to which common genetic variants associated with childhood cognition *en masse* explain variation in the genetic liability to schizophrenia, and (3) the extent to which the phenotypic correlation between cognition and schizophrenia is explained by common genetic effects.

## Methods

### Samples

**Cognition Samples.** The Avon Longitudinal Study of Parents and Children (ALSPAC) is a population-based birth cohort (<http://www.alspac.bris.ac.uk>) of, initially, 14 062 live births, and has been reported extensively.<sup>17,18</sup>

Phenotype data were collected by ALSPAC when participants were 8 years old. Individuals were administered the short form Wechsler Intelligence Scale for Children (WISC-III)<sup>19</sup> (alternate items used for all subtests except the coding subtest) as well as 2 additional items taken from the Test of Everyday Attention for Children (TEACH)<sup>20</sup> and the Diagnostic Analysis of Nonverbal Accuracy (DANVA).<sup>21</sup> Cognitive tests were selected a priori based on their correspondence to MATRICS domains as we have previously reported.<sup>22</sup> The following tasks were selected from the WISC-III, with their representative cognitive constructs in parentheses: coding (processing speed), digit span backward (working memory), and block design (problem solving). As a measure of verbal learning, we used the total number of nonwords correctly recalled from an adapted version of the Nonword Repetition Test.<sup>23</sup> From the TEACH, the Sky Search task adjusted for motor speed was selected as our measure of attention. The DANVA was used to create a social cognition variable by totaling number of errors (incorrect assignment of emotions) across all 4 emotional domains. The selection of cognitive tests and corresponding domains mirrored previous work using these samples.<sup>22</sup> We also used WISC IQ scores for verbal, performance and full IQ. The number of individuals with genotype and phenotype information in ALSPAC varied by specific test and ranged from 5109 to 5556 ([supplementary table 2](#)).

Further details on the ALSPAC sample including descriptive information about the cognitive variables, genotyping and imputation procedures can be found in [supplementary text](#) and [supplementary tables 1 and 2](#).

**Schizophrenia Samples.** Two independent schizophrenia samples were used as training datasets: the Schizophrenia Working Group of the Psychiatric Genetics Consortium (PGC-SCZ) and CLOZUK. The PGC-SCZ is a

collaborative effort across multiple centers with the aim of using large sample sizes to uncover novel findings in genetic analyses. The majority of samples in PGC-SCZ have a European ancestry. Information regarding individual samples and QC can be found elsewhere.<sup>16</sup> We use data from the CLOZUK sample as a replication dataset and therefore these individuals (cases and controls) were omitted from the PGC-SCZ sample for our primary analyses.

We further used CLOZUK as the testing dataset in the polygenic score analysis assessing the extent to which common genetic variants associated with childhood cognition explain variation in the genetic liability to schizophrenia, and in the bivariate GCTA. It was not possible to use the PGC-SCZ sample for this analysis given the lack of availability of genotype data and the requirement for this data in these analyses.

CLOZUK constitutes individuals with schizophrenia who were prescribed clozapine and received a clinician diagnosis of treatment resistant of schizophrenia. It has been demonstrated through a variety of different analyses that genetic risk factors for this sample are very similar to conventionally ascertained schizophrenia samples, including their frequency of CNVs,<sup>24</sup> sign test for schizophrenia associated loci,<sup>25</sup> and amount of common polygenic risk which aligns with other schizophrenia samples.<sup>16</sup> Therefore the genetic risk factors and patterns of association indicate this is a highly valid schizophrenia sample for these types of genetic interrogations.

GWAS summary statistics were available for the complete PGC-SCZ sample minus individuals from the CLOZUK sample (termed PGC-SCZ hereafter), totaling 29 415 schizophrenia cases and 40 101 controls. The CLOZUK sample consists of 5554 cases and 6299 controls and has been described elsewhere.<sup>16</sup>

Additional information about the schizophrenia samples and genotyping can be found in [supplementary text](#).

### Statistical Analysis

**Polygenic Score Analyses.** Polygenic score methodology follows the procedure described by the schizophrenia PGC.<sup>26</sup> For all analyses, SNPs were removed if they had a low minor allele frequency (MAF < 0.01), and were subsequently pruned for linkage disequilibrium using the clumping function (--clump) in PLINK<sup>27</sup> removing SNPs within 500kb (--clump-kb) and  $r^2 > .25$  (--clump-r2) with a more significantly associated SNP. We used the --score command in PLINK to calculate polygenic scores.<sup>27</sup> Five  $P$ -value thresholds ( $P_T < .0001, .01, .1, .3, .5$ ) for SNP association were used.

**Polygenic Risk Scores for Schizophrenia.** Genotypes were tested for association with schizophrenia in the 2 independent schizophrenia case/controls samples (1) PGC-SCZ and (2) CLOZUK.<sup>16</sup> Polygenic scores were

calculated by summing the number of susceptibility alleles of each index SNP weighted by the logarithm of the SNP odds ratios. Polygenic risk scores for schizophrenia were calculated for each individual in the ALSPAC sample and linear regressions were used to test whether these scores were associated with performance on cognitive tasks. Consistent with previous analyses using the ALSPAC sample, no covariates were included in the polygenic analysis as ALSPAC has previously been shown to have no significant population stratification.<sup>28</sup> We compared our original results against the results of 10 000 permutations where cognitive scores in ALSPAC individuals were randomly shuffled ([supplementary text](#)).

**Polygenic Risk Scores for Cognition.** Genotypes were tested for association with the cognitive phenotypes in ALSPAC using linear regression under an additive genetic model in PLINK.<sup>27</sup> No covariates were used in the analysis, as an EIGENSTRAT analyses did not reveal any significant population stratification. Analyses were run using the --standard-beta command in PLINK. Polygenic scores were calculated by summing the number of susceptibility alleles of each index SNP weighted by its coefficient of effect. Since genotype information is required in the target dataset, we used CLOZUK data for this analysis. Polygenic scores for each cognitive domain were calculated for each individual in the CLOZUK sample and logistic regressions were used to test whether the resulting scores were associated with schizophrenia case status. Analyses were covaried for 7 population principal components for the CLOZUK sample to ensure that estimates were not affected by subtle differences in population structure. We applied post hoc permutation correction for associations reaching nominal levels of significance. When applicable, we randomized case/control status in CLOZUK and repeated logistic regression analyses 10 000 times ([supplementary text](#)).

**Bivariate GCTA.** Bivariate GCTA was used to estimate the genetic correlation between schizophrenia risk and cognitive ability by common SNPs across the genome.<sup>14</sup> GCTA determines the degree of genetic relatedness between all individuals and then estimates the proportion of phenotypic variance explained by the genetic relatedness matrix using a mixed-effect model.<sup>13</sup> The bivariate extension to this approach determines the genetic variance of each trait as well as the genetic covariance, and then calculates the genetic correlation (the ratio between genetic covariance and genetic variance). We applied a stringent cut-off for relatedness (proportion of shared genetic material <2.5%) and removed 1 individual from each genetically related pair.<sup>13</sup> Bivariate GCTA was fitted with restricted maximum likelihood (REML) based on a linear model for the cognitive phenotypes and a disorder-model for schizophrenia with lifetime disorder risk set to 1%.<sup>29</sup> A genetic relationship matrix was created for CLOZUK

and ALSPAC. The genetic relationship matrix was analyzed for the genetic correlation ( $r_G$ ) between SNPs affecting schizophrenia case status and cognitive performance on the 6 domain tests and 3 IQ measures. The proportion of genetic variance ( $r_G^2$ ) shared between schizophrenia and each cognitive test was calculated.<sup>7,9</sup> As GCTA assesses the genetic relationship between individuals, a more rigorous correction for population stratification is warranted, therefore, 20 principal components were used as covariates.

## Results

### Polygenic Risk Scores for Schizophrenia: Association With Cognition

We examined if polygenic scores for schizophrenia were associated with cognitive performance, with the hypothesis that the schizophrenia risk alleles *en masse* would be associated with decreased cognitive performance. Associations with performance, verbal and full IQ are presented in **table 1** and **figure 1**, and associations with cognitive sub-domains are presented in **table 2** and **figure 1**. Schizophrenia polygenic risk scores derived from PGC-SCZ were associated with lower performance IQ in ALSPAC at age 8 ( $P = .001$ ;  $R^2 = 0.19\%$ ;  $P_T = .5$ ). This result was observed at multiple training set  $P$ -value thresholds and replicated using schizophrenia polygenic risk scores derived from the independent CLOZUK sample ( $P = .029$ ;  $R^2 = 0.09\%$ ;  $P_T = .5$ ).

Evidence of association with other cognitive tests was less consistent; eg, increased schizophrenia polygenic risk

was associated with lower full IQ when training on PGC-SCZ ( $P = .013$ ;  $R^2 = 0.11\%$ ;  $P_T = .5$ ) though not when training on CLOZUK. Permuted results were entirely consistent with the original regression results ([supplementary table 7](#)).

### Polygenic Risk Scores for Cognition: Association With Schizophrenia Status

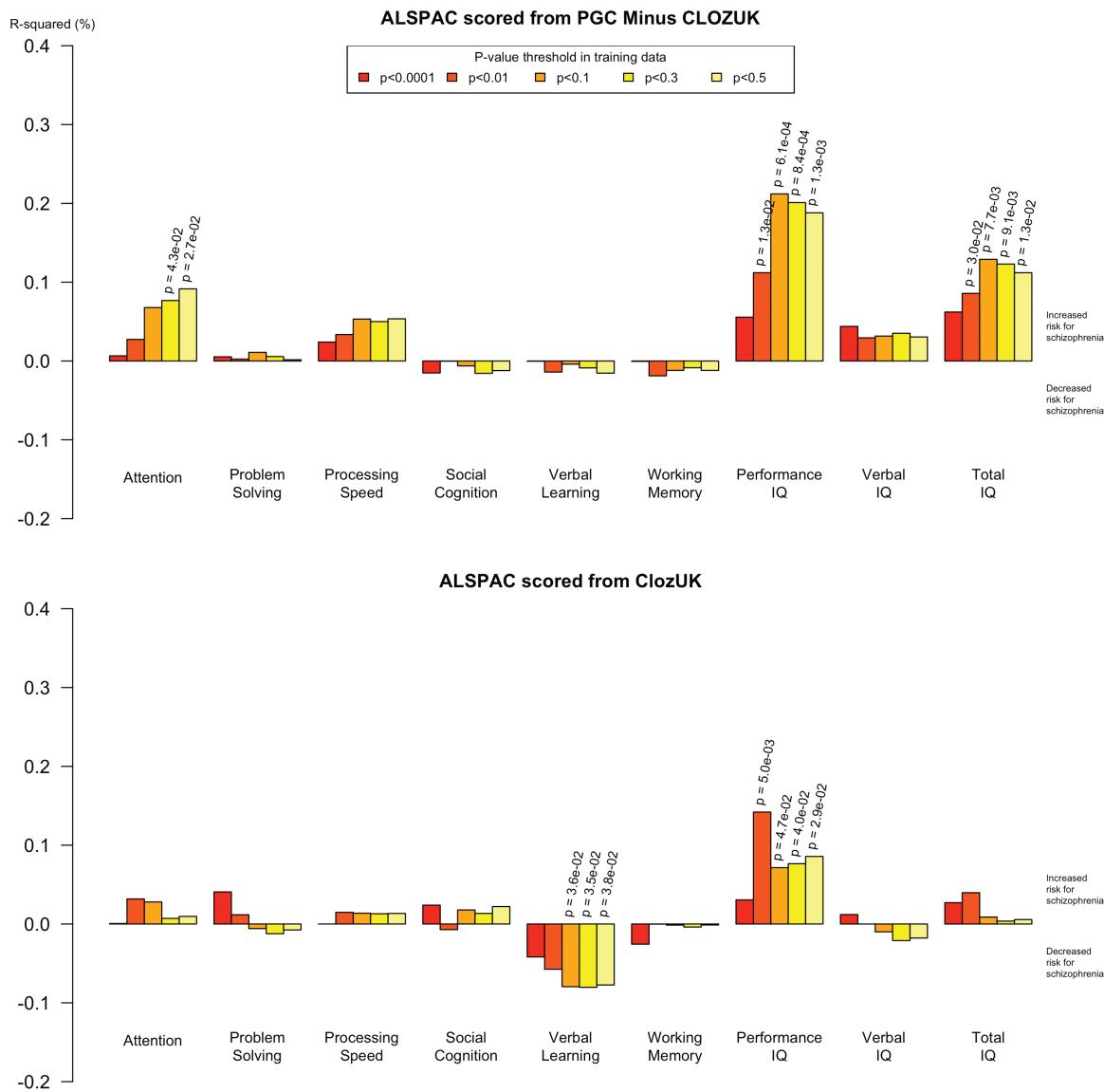
We investigated whether genetic variants for cognition at age 8 were associated with schizophrenia, with the hypothesis that genetic variants associated with lower cognitive performance would be associated with increased risk for schizophrenia. Performance IQ polygenic score derived from ALSPAC was significantly associated with schizophrenia case status in CLOZUK ( $P = 3.56E-04$ ;  $R^2 = 0.08\%$ ;  $P_T = .5$ ; **table 1** and **figure 2**). This association was seen at all training  $P$ -value thresholds. The direction of effect was consistent with our hypothesis. While there was limited evidence of association with processing speed, attention and verbal learning (**table 2** and **figure 2**), these associations were weaker and less consistent across training  $P$ -value thresholds than those for performance IQ. Permuted results were entirely consistent with the original regression results ([supplementary table 7](#)).

To provide results for comparison, we performed split half polygenic analysis for cognition within ALSPAC and for schizophrenia within the CLOZUK. Furthermore, work completed by the Schizophrenia Working Group of Psychiatric Genomics Consortium showed that polygenic

**Table 1.** Results From Polygenic Scores Analysis for the Global Measures of IQ

Training Dataset	Testing Dataset	$P_T$	N SNPs	Performance IQ			Verbal IQ			Full IQ		
				D	$R^2$ (%)	P-value	D	$R^2$ (%)	P-value	D	$R^2$ (%)	P-value
PGC-SCZ without cognition	ALSPAC	.0001	1497	–	0.0556	.079	–	0.0439	.119	–	0.0621	.064
		.01	16 721	–	0.112	.013	–	0.0294	.202	–	0.0858	.030
		.1	67 032	–	0.212	<i>6.07E-04</i>	–	0.0316	.186	–	0.129	.008
		.3	130 553	–	0.201	<i>8.43E-04</i>	–	0.0352	.163	–	0.123	.009
		.5	173 932	–	0.188	.001	–	0.0305	.194	–	0.112	.013
CLOZUK SCZ	ALSPAC cognition	.0001	372	–	0.0306	.193	–	0.026	.416	–	0.011	.221
		.01	10 133	–	0.142	.005	+	0.001	.965	–	0.022	.139
		.1	58 336	–	0.0715	.047	+	0.021	.458	–	0.007	.486
		.3	127 606	–	0.0765	.040	+	0.022	.281	–	0.005	.653
		.5	177 080	–	0.0857	.029	+	0.02	.321	–	0.005	.578
ALSPAC cognition	CLOZUK SCZ	.0001	32	–	0.028	.041	+	0.003	.727	–	0.008	.359
		.01	2157	–	0.042	.011	+	0.013	.200	–	0.014	.178
		.1	19 870	–	0.09	<i>1.56E-04</i>	+	0.004	.632	–	0.025	.058
		.3	50 819	–	0.084	<i>2.70E-04</i>	+	0.004	.611	–	0.016	.140
		.5	76 137	–	0.081	<i>3.56E-04</i>	+	0.005	.573	–	0.013	.198

*Note:* Italicized results are for  $P$ -value  $< .05$ . ALSPAC, Avon Longitudinal Study of Parents and Children; PGC-SCZ, Schizophrenia Working Group of the Psychiatric Genetics Consortium; SNP, single nucleotide polymorphism. Training dataset refers to the dataset used to create the polygenic scores and testing dataset is the set tested for prediction.  $P_T$  refers to the  $P$ -value threshold used in the training dataset. D is the direction of effect, with negative (–) meaning increased SCZ risk is associated with decreased cognitive performance and positive (+) indicating increased SCZ risk is associated with increased cognitive performance. The number of SNPs used in the scores for the ALSPAC cognitive to CLOZUK SCZ analyses varied slightly for each cognitive domain, here we reported the average but these numbers are reported in full in [supplementary table 4](#).



**Fig. 1.** Proportion of variance in cognitive phenotypes in the Avon Longitudinal Study of Parents and Children (ALSPAC) dataset explained by schizophrenia polygenic scores. The upper panel plot uses scores derived from the Schizophrenia Working Group of the Psychiatric Genetics Consortium (PGC-SCZ) with CLOZUK omitted (PGC – CLOZUK). The lower panel plot uses scores derived from the CLOZUK sample. Two-sided  $P$ -values for evidence at  $P < .05$  are displayed.  $R^2$  values above 0 reflect direction of effects consistent with our hypotheses that schizophrenia polygenic risk is associated with lower cognitive scores;  $R^2$  values below 0 reflect direction of effects opposite to those hypothesized.

risk scores using schizophrenia risk variants derived from PGC minus CLOZUK (PGC-SCZ) reliably distinguishes between schizophrenia cases and controls in CLOZUK.<sup>16</sup> At a  $P_{T} = .05$  the variance explained by the PGC-SCZ polygenic score was  $R^2 = 0.170$  in the CLOZUK sample ( $P$ -value  $< 4.26E-269$  [ $P$ -value reported in article is 0, therefore we report it as less than the last nonzero  $P$ -value reported]). Further details and results are presented in [supplementary tables 5 and 8](#) respectively.

#### Bivariate GCTA

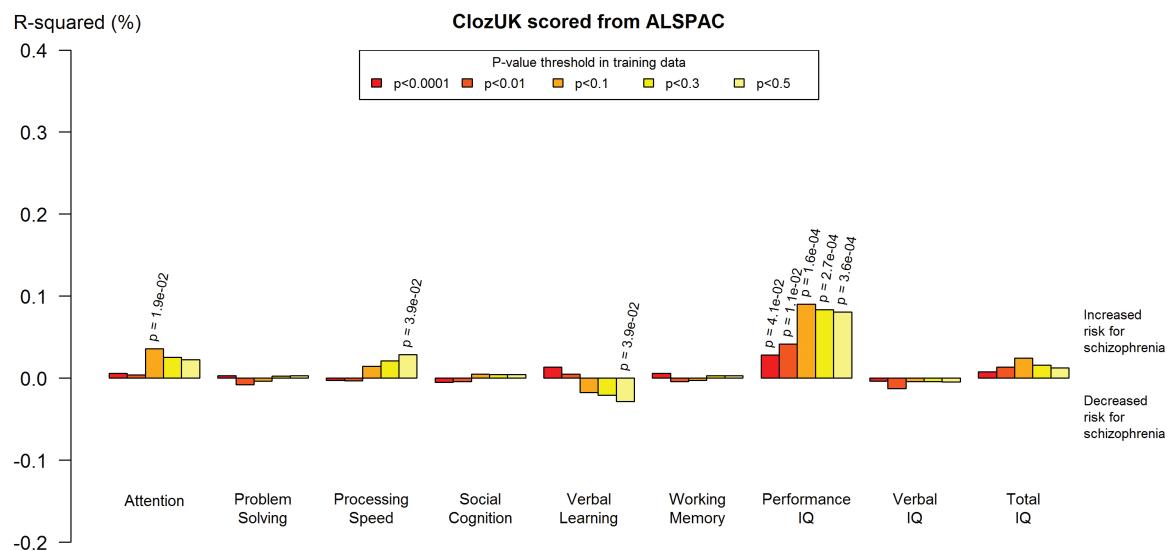
Bivariate GCTA revealed evidence of a common genetic overlap between schizophrenia and cognition. The

genetic correlation ( $r_G$ ) was  $-.379$  ( $P = 6.62E-05$ ) between performance IQ and schizophrenia, and  $-0.202$  between full IQ and schizophrenia ( $P = 5.00E-03$ ; [table 3](#)). The negative genetic correlation indicates that alleles associated with increased schizophrenia risk are associated with poorer cognitive function. Thus approximately 14% of the genetic component for schizophrenia is shared with that for performance IQ measured at age 8 and 4% for full IQ (estimated as  $r_G^2$ ). Evidence of genetic sharing was weaker for working memory and schizophrenia ( $r_G = -.193$ ;  $P = .044$ ), whilst there was no evidence to support the presence of shared genetic aetiology with schizophrenia for any of the other cognitive domains

**Table 2.** Results From Polygenic Scores Analysis for the Cognitive Sub-domains

Training Dataset	Testing Dataset	$P_T$	Attention			Problem Solving			Processing Speed			Social Cognition			Verbal Learning			Working Memory			
			N SNPs	D	$R^2(\%)$	P-value	D	$R^2(\%)$	P-value	D	$R^2(\%)$	P-value	D	$R^2(\%)$	P-value	D	$R^2(\%)$	P-value	D	$R^2(\%)$	P-value
PGC-SCZ without cognition	ALSPAC	.0001	1497	—	0.007	.555	—	0.005	.588	—	0.024	.248	+	0.015	.378	+	<0.001	.901	+	<0.001	.870
	CLOZUK	.01	16 721	—	0.027	.228	—	0.002	.716	—	0.034	.172	+	<0.001	.910	+	0.014	.379	+	0.019	.312
CLOZUK	ALSPAC	.1	67 032	—	0.068	.058	—	0.011	.436	—	0.053	.086	+	0.006	.573	+	0.004	.636	+	0.012	.423
	SCZ cognition	.3	130 553	—	0.077	.043	—	0.006	.577	—	0.050	.095	+	0.016	.370	+	0.009	.486	+	0.009	.495
ALSPAC	ALSPAC	.5	173 932	—	0.092	.027	—	0.002	.767	—	0.054	.085	+	0.012	.433	+	0.015	.355	+	0.012	.422
	CLOZUK	.0001	372	—	0.001	.856	—	0.041	.135	+	<0.001	.983	—	0.024	.269	+	0.042	.128	+	0.026	.239
ALSPAC cognition	ALSPAC	.01	10 133	—	0.032	.194	—	0.012	.422	—	0.015	.365	+	0.007	.547	+	0.057	.074	+	<0.001	.933
	CLOZUK	.1	58 336	—	0.028	.222	+	0.006	.571	—	0.014	.384	—	0.018	.340	+	0.080	.036	+	0.001	.782
ALSPAC cognition	ALSPAC	.3	127 606	—	0.007	.538	+	0.012	.411	—	0.013	.396	—	0.013	.407	+	0.080	.035	+	0.004	.648
	CLOZUK	.5	177 080	—	0.010	.474	+	0.008	.519	—	0.013	.389	—	0.022	.288	+	0.077	.038	+	0.001	.795
CLOZUK SCZ	ALSPAC	.0001	32	—	0.006	.460	—	0.003	.820	+	0.003	.897	+	0.005	.523	—	0.014	.180	—	0.006	.455
	SCZ	.01	21 57	—	0.004	.635	+	0.008	.354	+	0.003	.827	+	0.004	.609	—	0.005	.556	+	0.004	.607
ALSPAC cognition	ALSPAC	.1	19 870	—	0.036	.019	+	0.004	.708	—	0.014	.165	—	0.005	.517	+	0.017	.120	+	0.003	.979
	CLOZUK	.3	50 819	—	0.026	.052	—	0.003	.908	—	0.021	.081	—	0.004	.604	+	0.021	.084	—	0.003	.777
CLOZUK SCZ	ALSPAC	.5	76 137	—	0.023	.070	—	0.003	.766	—	0.029	.039	—	0.005	.573	+	0.029	.039	—	0.003	.842

*Note:* Italicized results are for  $P$ -value  $< .05$ . ALSPAC, Avon Longitudinal Study of Parents and Children; PGC-SCZ, Schizophrenia Working Group of the Psychiatric Genetics Consortium; SNP, single nucleotide polymorphism. Training dataset refers to the dataset used to create the polygenic scores and testing dataset is the set tested for prediction.  $P_T$  refers to the  $P$ -value threshold used in the training dataset. D is the direction of effect, with negative (—) meaning increased SCZ risk is associated with decreased cognitive performance and positive (+) indicating increased SCZ risk is associated with increased cognitive performance. The number of SNPs used in the scores for the ALSPAC cognitive to CLOZUK SCZ analyses varied slightly for each cognitive domain, here we reported the average but these numbers are reported in full in [supplementary table 4](#).



**Fig. 2.** Proportion of variance in schizophrenia explained by polygenic scores of alleles associated with different cognitive phenotypes within the Avon Longitudinal Study of Parents and Children (ALSPAC) sample. The figure plots polygenic scores derived from ALSPAC cognitive phenotypes tested in the CLOZUK schizophrenia sample. Two-sided  $P$ -values for evidence at  $P < .05$  are displayed.  $R^2$  values above 0 reflect direction of effects consistent with our hypotheses that polygenic risk for lower cognition scoring is associated with higher schizophrenia risk;  $R^2$  values below 0 reflect direction of effects opposite to those hypothesized.

(table 3). Univariate GCTA results for each phenotype can be found in supplementary table 6.

## Discussion

We have investigated the relationship between cognition and schizophrenia using data from a large population sample and multiple independent, large schizophrenia datasets. We tested both whether polygenic scores for cognition measured at age 8 were associated with schizophrenia and if SNPs conferring risk for schizophrenia were associated with cognitive performance in children at age 8. We found a consistent, reciprocal relationship between SNPs associated with schizophrenia and SNPs associated with performance IQ using polygene scoring analysis and bivariate GCTA. There was no evidence to support genetic overlap between schizophrenia and verbal IQ, and hence weaker evidence for full IQ (as this is constructed from performance and verbal IQ). The finding that the genetic overlap between cognition and schizophrenia is strongest for performance IQ is noteworthy and should inform endophenotype research strategies particularly given previous findings that performance IQ is compromised when compared to verbal IQ in patients and their relatives.<sup>30,31</sup> Little evidence was found to support genetic overlap between schizophrenia and other specific cognitive domains using either the polygenic score analysis or bivariate GCTA.

In order to quantify the genetic overlap between schizophrenia and cognition, we used the genetic correlation ( $r_G$ ), which indexes the correlation between genetic effects of 2 phenotypes. The genetic correlation between full IQ and schizophrenia that we observed

using bivariate GCTA ( $r_G = -.202$ ) was smaller than estimated through 2 twin studies that identified twin pairs on the basis of their case status and measurements of cognition post onset of schizophrenia ( $r_G = -.75$  and  $r_G = -.46$ , respectively). It should be noted that our estimates are based only on variants tagged by common SNPs<sup>13</sup> and that these comparisons are limited by differences in cognitive measures used, sample ascertainment strategies, and differing study designs. Despite this our findings align well with a population-based twin study that reported a genetic correlation between psychosis and premorbidly-assessed full IQ of  $-0.26$ .<sup>9</sup> The similar association between IQ and schizophrenia across different co-relative groups reported in a recent population-based study also suggests that shared genetic effects are likely to be very modest.<sup>32</sup>

Our results partially replicate previous work investigating the genetic overlap between schizophrenia and cognition using polygenic score analysis that focused solely on full IQ or on “g.”<sup>11,12</sup> The amount of variance in full IQ explained by schizophrenia polygenic risk in our study ( $R^2 \sim 0.1\%$ ) is lower than reported for “g” ( $R^2 \sim 0.5\%$  to 1%) even though our training dataset was substantially larger.<sup>11</sup> Our study was equivalent in terms of the size of the cognitive datasets, but all studies to date are likely to have been underpowered to accurately measure genetic effects on cognition,<sup>12</sup> which might explain why results across studies are inconsistent in estimating effect sizes.<sup>33</sup> It is also possible that our study differs from previous studies in terms of differences in the cognitive tests used or age at which the cognitive tests were assessed. Nevertheless, it is apparent that the magnitude of the overlapping polygenic signal between schizophrenia and

**Table 3.** Results From the Bivariate GCTA Analysis

Schizophrenia Dataset	Cognitive Tests From ALSPAC	Schizophrenia # (Case/Control)	ALSPAC #	# SNPs	$h_{\text{SNP}}^2$ (SE)	$h_{\text{SNP}}^2$ (SE) Cognitive Test	$r_G$ (SE)	$r_G^2$	P-value
CLOZUK									
Attention	5336/6023	4711	424 126	0.368 (0.016)	0.006 (0.083)	1.00 (6.845)	1	.116	
Problem solving	5336/6023	4869	424 126	0.369 (0.016)	0.286 (0.081)	−0.094 (0.086)	.009	.268	
Processing speed	5336/6023	4919	424 126	0.370 (0.016)	0.238 (0.082)	−0.147 (0.096)	.022	.114	
Social cognition	5336/6023	4519	424 126	0.369 (0.016)	<0.001 (0.087)	1.00 (174.306)	1	.736	
<b>Working memory</b>	<b>5336/6023</b>	<b>4797</b>	<b>424 126</b>	<b>0.370 (0.016)</b>	<b>0.231 (0.083)</b>	<b>−.193 (0.101)</b>	<b>.037</b>	<b>.044</b>	
Verbal learning	5336/6023	4916	424 126	0.369 (0.016)	0.278 (0.081)	0.036 (0.087)	.001	.680	
<b>Performance IQ</b>	<b>5336/6023</b>	<b>4902</b>	<b>424 126</b>	<b>0.369 (0.016)</b>	<b>0.231 (0.080)</b>	<b>−.379 (0.113)</b>	<b>.144</b>	<b>6.62E-05</b>	
Verbal IQ	5336/6023	4905	424 126	0.370 (0.016)	0.511 (0.080)	−0.067 (0.064)	.004	.296	
<b>Full IQ</b>	<b>5336/6023</b>	<b>4886</b>	<b>424 126</b>	<b>0.370 (0.016)</b>	<b>0.403 (0.080)</b>	<b>−.202 (0.074)</b>	<b>.041</b>	<b>5.00E-03</b>	

*Note:* Bold and italicized results are for  $P$ -value  $< .05$ .  $h_{\text{SNP}}^2$ , SNP based heritability;  $r_G$ , genetic correlation; SNP, single nucleotide polymorphism. Two-sided  $P$ -values reported testing if  $r_G$  is different from 0.

cognition is modest, as measured by different studies to date.

As heritability of IQ increases throughout childhood and early adolescence,<sup>34,35</sup> it may be that genetic effects are contributing to cognition to a lesser extent in our study, which assessed cognition during childhood, than in previous studies that assessed cognition during adulthood. However, GCTA estimates of common genetic factors contributing to IQ in childhood are still substantial ( $h_{\text{SNP}}^2$  ranging from 0.22 to 0.46)<sup>36</sup> and we observe similar estimates in the present study (table 3 for  $h_{\text{SNP}}^2$  for cognitive test in ALSPAC). Furthermore, these estimates are similar to those from twin studies ( $h^2 = 0.41$ ).<sup>37</sup>

### Implications

Our comprehensive approach to investigating common genetic architecture across cognition and schizophrenia leads us to 3 important observations. First, whilst genetic effects are shared between cognition measured at age 8 and schizophrenia this is primarily with performance IQ, implicating fluid rather than crystallized intelligence.<sup>38</sup> Despite evidence of shared genetic effects between performance IQ and schizophrenia, the extent to which common SNPs explain variation in these phenotypes is modest. While there was a sizeable genetic correlation between schizophrenia and performance IQ ( $r_G = -.38 \pm 0.11$  SE), this is smaller (though less precisely estimated) than that reported between schizophrenia and major depressive disorder ( $r_G = .43 \pm 0.06$  SE), and substantially smaller than between schizophrenia and bipolar disorder ( $r_G = .68 \pm 0.04$  SE), although higher than several other neuropsychiatric disorders.<sup>15</sup>

Second, most cognitive measures that we examined, which map onto the MATRICS domains, show little or no evidence of genetic overlap with schizophrenia. If this is replicated in other studies, it will call into question their use as endophenotypes. Thirdly the fact that cognition measured in a general population sample at age 8 shows genetic overlap with schizophrenia provides support for a genetic contribution to the neurodevelopmental nature of schizophrenia.

Endophenotypes as originally conceived were seen as aiding gene discovery for phenotypes such as schizophrenia. Our results, together with those in adult populations, suggest cognitive measures will have limited utility in this sense. A more productive approach may be to leverage the genetic overlap between cognition and schizophrenia by using cognition as a stratifier in schizophrenia discovery GWAS.<sup>39</sup>

It is possible that our findings of limited genetic sharing reflect the fact that we have only examined the effects of single common genotyped variants that collectively explain limited heritability or that measures used to date do not adequately capture aspects of cognition that are most relevant to the genetic aetiology of schizophrenia.

For example, cognitive models of psychosis suggest that deficits in source monitoring, reasoning biases and predictive error processing may play key roles in the aetiology of psychotic experiences, and provide an explanatory link between altered biological processes and psychopathology.<sup>40–42</sup> Whilst numerous studies have measures of cognitive ability related to IQ, working memory, processing speed and attention, very few studies have measures that directly capture deficits in the cognitive models of psychosis described above. If these show greater merit as endophenotypes for schizophrenia than currently assessed cognitive domains, their value will most likely be in helping to understand the mechanisms mediating genetic risk for schizophrenia as identified through large-scale collaborative genome-wide enterprises (eg, recent genome-wide evidence of association for over 100 SNPs for schizophrenia from a consortium of >35 000 cases).<sup>16</sup>

### Strengths and Limitations

Our study has many strengths including use of the world's largest schizophrenia dataset, a large and relatively homogenous population based sample for assessing cognition, and use of cognitive sub-domains as well as general IQ measures. Furthermore, we assess the genetic relationship between schizophrenia and cognition using 2 different complementary approaches, polygenic analysis and bivariate GCTA. One of the main limitations of our study is that despite the use of one of the largest population-based birth cohorts worldwide with the required cognitive and genetic data, this sample may be too small to serve as a training dataset for deriving polygenic scores of cognitive phenotypes. However, to our knowledge, this is the largest homogenous collection of childhood cognitive data available currently. Furthermore, we have used only cognitive data measured during childhood and are not able to examine the possible influence of varying heritability at different ages, although our results using GCTA provide evidence that cognition at age 8 clearly has a substantial common genetic component. It is challenging to make comparisons between variance explained in "within trait" analyses (such as that presented in [supplementary tables 5 and 8](#)) and the primary cross-trait analyses we present between schizophrenia and cognition, particularly given the difference in power between samples. Such comparisons will be aided by equivalent studies conducted in very large cognitive datasets in which the phenotypic variance explained by common GWAS variants reaches a plateau as has begun to be observed in the PGC-SCZ datasets.<sup>16</sup> A potential limitation is that we cannot identify individuals in the ALSPAC sample that will develop schizophrenia, and that these individuals may bias the results. However, this would apply to less than 1% of the ALSPAC sample and given the normal distribution of polygenic scores for schizophrenia and cognition it seems inconceivable that they would explain the observed

results. Another limitation is that rare genetic variants are not captured by GWAS, and therefore do not contribute to our estimates; we are therefore capturing the lower bound estimate of the genetic overlap between cognition and schizophrenia. Individuals with lower cognition were also more likely to have missing genetic data ([supplementary table 3](#)), thus our estimates of the genetic overlap between schizophrenia and cognition is likely to be conservative. A final limitation is that we conducted tests at multiple *P*-value thresholds across cognitive outcomes. Whilst there is no established methodology for correcting for such correlated polygenic tests we did confirm nominally significant results with permutation. Significant associations between schizophrenia and performance IQ were observed in both polygenic score and GCTA analyses and we feel this is unlikely to have arisen due to type I error resulting from multiple testing.

In conclusion, we have undertaken a detailed examination of the common genetic architecture across cognition and schizophrenia. Our results show the relationship between schizophrenia and performance IQ is in part due to shared common genetic factors, and that this relationship is evident using a childhood cognitive sample. Our results suggest that identifying genetic factors underlying cognitive endophenotypes will capture only modest proportions of schizophrenia genetic risk. The true value of cognitive endophenotypes may lie in helping to understand the mechanisms mediating the relationship between genes and disorder.

### Supplementary Material

Supplementary material is available at <http://schizophreniabulletin.oxfordjournals.org>.

### Funding

UK Medical Research Council (74882); the Wellcome Trust (076467); UK Medical Research Council (MR/K004360/1 and MR/M006727/1 to D.E.J.L. and paid for K.E.T.'s salary); University of Bristol provided core support for ALSPAC. Work at the Broad Institute was funded by a philanthropic gift to the Stanley Center for Psychiatric Research. L.H. was supported by a UK Medical Research Council PhD studentship.

### Acknowledgments

We are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, including interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. ALSPAC GWAS data was generated by Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute and LabCorp (Laboratory Corporation of America) using support

from 23andMe. For the CLOZUK sample we thank Novartis for their guidance and cooperation. We also thank staff at The Doctor's Laboratory, UK for help and advice regarding sample acquisition. We also acknowledge the Schizophrenia Working Group of the Psychiatric Genomics Consortium for publicly available data used to derive schizophrenia polygenic risk scores. The authors declare no conflicts of interest.

## References

1. Green MF. Cognitive impairment and functional outcome in schizophrenia and bipolar disorder. *J Clin Psychiatry*. 2006;67:3–8; discussion 36–42.
2. Reichenberg A, Caspi A, Harrington H, et al. Static and dynamic cognitive deficits in childhood preceding adult schizophrenia: a 30-year study. *Am J Psychiatry*. 2010;167:160–169.
3. Zammit S, Allebeck P, David AS, et al. A longitudinal study of premorbid IQ Score and risk of developing schizophrenia, bipolar disorder, severe depression, and other nonaffective psychoses. *Arch Gen Psychiatry*. 2004;61:354–360.
4. Woodberry KA, Giuliano AJ, Seidman LJ. Premorbid IQ in schizophrenia: a meta-analytic review. *Am J Psychiatry*. 2008;165:579–587.
5. Sponheim SR, Jung RE, Seidman LJ, et al. Cognitive deficits in recent-onset and chronic schizophrenia. *J Psychiatr Res*. 2010;44:421–428.
6. Nuechterlein KH, Barch DM, Gold JM, Goldberg TE, Green MF, Heaton RK. Identification of separable cognitive factors in schizophrenia. *Schizophr Res*. 2004;72:29–39.
7. Toulopoulou T, Picchioni M, Rijsdijk F, et al. Substantial genetic overlap between neurocognition and schizophrenia: genetic modeling in twin samples. *Arch Gen Psychiatry*. 2007;64:1348–1355.
8. Toulopoulou T, Goldberg TE, Mesa IR, et al. Impaired intellect and memory: a missing link between genetic risk and schizophrenia? *Arch Gen Psychiatry*. 2010;67:905–913.
9. Fowler T, Zammit S, Owen MJ, Rasmussen F. A population-based study of shared genetic variation between premorbid IQ and psychosis among male twin pairs and sibling pairs from Sweden. *Arch Gen Psychiatry*. 2012;69:460–466.
10. McIntosh AM, Gow A, Luciano M, et al. Polygenic risk for schizophrenia is associated with cognitive change between childhood and old age. *Biol Psychiatry*. 2013;73:938–943.
11. Lencz T, Knowles E, Davies G, et al. Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: a report from the Cognitive Genomics consorTium (COGENT). *Mol Psychiatry*. 2014;19:168–174.
12. van Scheltinga AF, Bakker SC, van Haren NE, et al. Schizophrenia genetic variants are not associated with intelligence. *Psychol Med*. 2013;43:2563–2570.
13. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88:76–82.
14. Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics*. 2012;28:2540–2542.
15. Cross-Disorder Group of the Psychiatric Genomics C, Lee SH, Ripke S, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet*. 2013;45:984–994.
16. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511:421–427.
17. Fraser A, Macdonald-Wallis C, Tilling K, et al. Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol*. 2013;42:97–110.
18. Boyd A, Golding J, Macleod J, et al. Cohort profile: the 'children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol*. 2013;42:111–127.
19. Wechsler D, Golombok S, Rust J. *WISC-III UK Wechsler Intelligence Scale for Children*. 3rd ed UK Manual. Sidcup, UK: The Psychological Corporation; 1992.
20. Robertson IH, Ward T, Ridgeway V, Nimmo-Smith I. The structure of normal human attention: the test of everyday attention. *J Int Neuropsychol Soc*. 1996;2:525–534.
21. Nowicki S, Duke M. Individual differences in the nonverbal communication of affect: the Diagnostic Analysis of NonVerbal Accuracy scale. *J Nonverbal Behav*. 1994;18:9–35.
22. Niarchou M, Zammit S, Walters J, Lewis G, Owen MJ, van den Bree MB. Defective processing speed and nonclinical psychotic experiences in children: longitudinal analyses in a large birth cohort. *Am J Psychiatry*. 2013;170:550–557.
23. Gathercole SE, Willis CS, Baddeley AD, Emslie H. The Children's Test of Nonword Repetition: a test of phonological working memory. *Memory*. 1994;2:103–127.
24. Rees E, Walters JT, Georgieva L, et al. Analysis of copy number variations at 15 schizophrenia-associated loci. *Br J Psychiatry*. 2014;204:108–114.
25. Hamshere ML, Walters JT, Smith R, et al. Genome-wide significant associations in schizophrenia to ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by the Schizophrenia PGC. *Mol Psychiatry*. 2013;18:708–712.
26. Schizophrenia PGC. Genome-wide association study identifies five new schizophrenia loci. *Nature Genet*. 2011;43:969–976.
27. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–575.
28. Martin J, Hamshere ML, Stergiakouli E, O'Donovan MC, Thapar A. Neurocognitive abilities in the general population and composite genetic risk scores for attention-deficit hyperactivity disorder. *J Child Psychol Psychiatry*. 2014;56:648–656.
29. McGrath J, Saha S, Chant D, Welham J. Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiol Rev*. 2008;30:67–76.
30. Heinrichs RW, Zakzanis KK. Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. *Neuropsychology*. 1998;12:426–445.
31. Kravariti E, Toulopoulou T, Mapua-Filbey F, et al. Intellectual asymmetry and genetic liability in first-degree relatives of probands with schizophrenia. *Br J Psychiatry*. 2006;188:186–187.
32. Kendler KS, Ohlsson H, Sundquist J, Sundquist K. IQ and schizophrenia in a Swedish national sample: their causal relationship and the interaction of IQ with genetic risk. *Am J Psychiatry*. 2015;172:259–265.
33. Button KS, Ioannidis JP, Mokrysz C, et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*. 2013;14:365–376.

34. Trzaskowski M, Yang J, Visscher PM, Plomin R. DNA evidence for strong genetic stability and increasing heritability of intelligence from age 7 to 12. *Mol Psychiatry*. 2014;19:380–384.
35. Bouchard TJ. The Wilson Effect: the increase in heritability of IQ with age. *Twin Res Hum Genet*. 2013;16:923–930.
36. Benyamin B, Pourcain B, Davis OS, et al. Childhood intelligence is heritable, highly polygenic and associated with FNBP1L. *Mol Psychiatry*. 2014;19:253–258.
37. Haworth CM, Wright MJ, Luciano M, et al. The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol Psychiatry*. 2010;15:1112–1120.
38. Christoforou A, Espeseth T, Davies G, et al. GWAS-based pathway analysis differentiates between fluid and crystallized intelligence. *Genes Brain Behav*. 2014;13:663–674.
39. Andreassen OA, Djurovic S, Thompson WK, et al. Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. *Am J Hum Genet*. 2013;92:197–209.
40. Garety P, Freeman D, Jolley S, Ross K, Waller H, Dunn G. Jumping to conclusions: the psychology of delusional reasoning. *Adv Psychiatr Treat*. 2011;17:332–339.
41. Waters F, Woodward T, Allen P, Aleman A, Sommer I. Self-recognition deficits in schizophrenia patients with auditory hallucinations: a meta-analysis of the literature. *Schizophr Bull*. 2012;38:741–750.
42. Fletcher PC, Frith CD. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci*. 2009;10:48–58.