

Higher genetic risk of schizophrenia is associated with lower cognitive performance in healthy individuals

Authors:

Rebecca Shafee^{1,2}, Pranav Nanda³, Jaya L. Padmanabhan⁴, Neeraj Tandon⁵, Ney Alliey-Rodriguez⁶, Richard S. E. Keefe⁷, Scot K. Hill⁸, Jeffrey R. Bishop⁹, Brett A. Clementz¹⁰, Carol A. Tamminga¹¹, Elliot S. Gershon^{6,12}, Godfrey D. Pearlson¹³, Matcheri S. Keshavan⁴, John A. Sweeney¹⁴, Elise B. Robinson^{2,15,16}, Steven A. McCarroll^{1,2,16}

¹Department of Genetics, Harvard Medical School, Boston, MA, USA

²Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA

³College of Physicians & Surgeons, Columbia University Medical Center, 710 W 168th St, New York, NY, USA

⁴Department of Psychiatry and Cognitive Neurology, Beth Israel Deaconess Medical Center, Boston, MA, USA

⁵Baylor College of Medicine, Texas Medical Center, Houston, TX, USA

⁶Department of Psychiatry and Behavioral Neuroscience, University of Chicago, IL, USA

⁷Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA

⁸Department of Psychology, Rosalind Franklin University of Medicine and Science, North Chicago, Illinois, USA

⁹Department of Pharmacy and Psychiatry, University of Minnesota, Minneapolis, MN, USA

¹⁰Department of Psychology and Neuroscience, University of Georgia, Athens, GA, USA

¹¹Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, Texas, USA

¹²Department of Human Genetics, University of Chicago, IL, USA

¹³Departments of Psychiatry and Neurobiology, Yale University and Olin Neuropsychiatric Research Center, Hartford, Connecticut, USA

¹⁴Department of Psychiatry and Behavioral Neuroscience, University of Cincinnati, OH, USA

¹⁵Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

¹⁶Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA

Corresponding Author: Rebecca Shafee, rebecca_shafee@hms.harvard.edu

Keywords: schizophrenia, genetic risk, cognition

Abstract

Psychotic disorders including schizophrenia are commonly accompanied by cognitive deficits. Recent studies have reported negative genetic correlations between schizophrenia and indicators of cognitive ability such as general intelligence and processing speed. Here we compare the effect of the genetic risk of schizophrenia (PRS_{SCZ}) on measures that differ in their relationships with psychosis onset: a measure of current cognitive abilities (the Brief Assessment of Cognition in Schizophrenia, BACS) that is greatly reduced in psychosis patients; a measure of premorbid intelligence that is minimally affected by psychosis (the Wide-Range Achievement Test, WRAT); and educational attainment (EY), which covaries with both BACS and WRAT. Using genome-wide SNP data from 314 psychotic and 423 healthy research participants in the Bipolar-Schizophrenia Network for Intermediate Phenotypes (B-SNIP) Consortium, we investigated the association of PRS_{SCZ} with BACS, WRAT and EY. Among apparently healthy individuals, greater genetic risk for schizophrenia (PRS_{SCZ}) was associated with lower BACS scores ($r = -0.19$, $p = 1 \times 10^{-4}$ at $P_T = 1 \times 10^{-4}$) but did not associate with WRAT or EY, suggesting that these areas of cognition vary in their etiologic relationships with schizophrenia. Among individuals with psychosis, PRS_{SCZ} did not associate with variation in cognitive performance. These findings suggest that the same cognitive abilities that are disrupted in psychotic disorders are also associated with schizophrenia genetic risk in the general population. Specific cognitive phenotypes, independent of education or general intelligence, could be more deeply studied for insight into the specific processes affected by the genetic influences on psychosis.

Significance

Psychotic disorders such as schizophrenia often involve profound cognitive deficits, the genetic underpinnings of which remain to be elucidated. Poor educational performance early in life is a well-known risk factor for future psychotic illness, potentially reflecting either shared genetic influences or other risk factors that are epidemiologically correlated. Here we show that, in apparently healthy individuals, common genetic risk factors for schizophrenia associate with lower performance in areas of cognition that are impaired in psychotic disorders but do not associate independently with educational attainment or more general measures of intelligence. These results suggest that specific cognitive phenotypes – independent of education or general intelligence – could be more deeply studied for insight into the processes affected by the genetic influences on psychosis.

Introduction

Schizophrenia is a debilitating psychiatric disorder that commonly involves severe cognitive deficits that compromise functional ability. Although at present schizophrenia is classified as a psychotic disorder, it has been suggested that it is primarily a cognitive disorder (1,2). Cognitive deficits have been reported not only in schizophrenia but also in other psychotic disorders (3,4). Underperformance in general intelligence tasks as well as tasks designed to be specific to cognitive domains such as memory, executive function, motor function etc. have been noted in psychosis patients (5).

Many of these cognitive deficits are present many years prior to the development of the illness (6,7). A meta-analysis of 4396 schizophrenia cases and 745,000 controls showed that every point decrease in premorbid IQ associates with a 3.7% increase in schizophrenia risk (8). In another analysis in a nationwide cohort of over 900,000 Swedish individuals, children with the lowest grades showed a 4-fold increased risk of developing schizophrenia and schizoaffective-disorder and a 3-fold increased risk of developing other psychotic illnesses (9). Additionally, studies of clinically high-risk (CHR) groups have shown that these near-psychotic individuals are cognitively impaired compared to healthy controls and that within the CHR group those that convert to psychosis in the future display lower cognitive performance compared to the nonconverters (10,11,12,13). There is also evidence of lower cognitive performance in first-episode psychosis patients compared to the CHR group (10,12,13). Together these results indicate that cognitive deficits are significantly associated with risk of developing a psychotic illness. It has also been reported that CHR individuals that convert to

psychosis show longitudinal decline in neurocognitive performance as they progress to acute psychosis (14). Although the relative roles played by cognitive deficit and cognitive decline along the psychosis onset trajectory are still not well understood, it is clear that lower cognitive function is related to increased vulnerability to psychotic illness onset.

Both cognitive performance and schizophrenia are heritable (15-23), and significant genetic overlap has been reported between schizophrenia and indicators of cognitive ability, such as general intelligence or processing speed (24-30). However, it is still unclear how the genetic differences associated with schizophrenia influence cognitive function, and which domains of cognitive function are most associated with schizophrenia risk.

Motivated by these earlier findings, we investigated the relationship of the genetic risk for schizophrenia – as defined by large constellations of common variants that associate with schizophrenia risk (“polygenic risk”, PRS_{SCZ}) – on three phenotypes: 1. the Brief Assessment of Cognition in Schizophrenia (BACS, 31), which consists of six subtests spanning multiple cognitive domains and provides a composite score of general cognitive function, 2. the Wide-Range Achievement Test (WRAT, 32-34) reading score, a measure of premorbid intellectual potential; and 3. educational attainment (as measured by years of education, EY), which is phenotypically associated with WRAT and BACS and also genetically overlaps with cognition (35,36). These analyses were done using the psychotic probands (PSYCH, $N = 314$) and nonpsychotic individuals (NPSYCH, $N = 423$) of the Bipolar-Schizophrenia Network of Intermediate Phenotypes consortium (B-SNIP, 3, 37). The PSYCH group included psychotic schizophrenia, bipolar disorder and

schizoaffective disorder probands and the NPSYCH group was formed by combining healthy controls and nonpsychotic family members (more details in Methods).

These measures of cognitive performance and educational achievement vary in their relationships to schizophrenia. While psychosis probands show lower values for all three phenotypes compared to healthy control individuals, the deficit in BACS in the patient group is much more significant than the deficit in EY or WRAT (3). In our sample the psychotic patient group's mean BACS performance was one standard deviation lower than the mean of the nonpsychotic group, whereas their EY and WRAT scores were lower by 0.3 standard deviations (Figures S1). Also, while many cognitive domains captured by BACS, such as, working memory and processing speed, show increasing deficits along the psychosis onset trajectory (decreasing performance from CHR individuals that do not convert to psychosis, to CHR individuals that do convert to psychosis to first-episode psychosis patients, 10-13), WRAT is a relatively stable measurement that is minimally affected by psychosis onset (10) and is commonly used as a measure for premorbid intelligence (32-34). Based on these phenotypic observations we hypothesized that the genetic risk of schizophrenia should be more strongly associated with BACS than with WRAT. Additionally, we investigated the effect of the polygenic predictors of educational attainment (PRS_{EDUC}) on BACS, EY and WRAT. We find that in healthy individuals, genetic risk for schizophrenia is associated with BACS, but not with WRAT or EY. This finding suggests that specific domains of cognition may be more closely etiologically linked to schizophrenia than other domains are, creating an opportunity for longitudinal studies to identify the domains that best predict illness onset.

Results

Validation of genetic data: genetic risk for schizophrenia is higher among psychosis patients

Polygenic liability for a trait – an estimate of genetic risk – is calculated by adding the relatively small effects of many individual common variants, as estimated from very large earlier genetic studies (the 2014 schizophrenia Genome-Wide Association Study by PGC consisted of 36,989 cases and 113,075 controls, 22). The score can be used to predict the value of a continuous trait (e.g. height) or the risk of a disease. For schizophrenia, polygenic scores predict 18.4% of the variation in disease risk (22). These scores are often calculated at multiple p-value thresholds (P_T). Lower p-value thresholds include fewer SNPs, each with more significant associations. Higher p-value thresholds include more SNPs, but the effect of the risk-conferring SNPs can be diluted among a larger fraction of neutral SNPs. This varying P_T approach is exploratory and used when the P_T for which the predictive power of the score is maximum for a particular phenotype is unknown.

An individual's PRS_{SCZ} represents common genetic influences on the risk of developing schizophrenia. Hence one would expect higher PRS_{SCZ} in the PSYCH group compared to the NPSYCH group. As a validation analysis of our genetic data we investigated the differences in PRS_{SCZ} between the PSYCH and NPSYCH groups. As shown in Figure 1, compared to the NPSYCH group the PSYCH group showed significantly higher PRS_{SCZ} at all P_T as expected ($p \leq P_{FDR} = 8.9 \times 10^{-4}$). Among the psychosis probands schizophrenia patients had highest PRS_{SCZ} (Figure S2). In our sample

PRS_{EDUC} did not differ significantly between the PSYCH and the NPSYCH groups at any P_T. Effect sizes and p-values for PSYCH/NPSYCH group differences in PRS_{SCZ} and PRS_{EDUC} can be found in Table S1. Figure S2 shows the distributions of PRS_{SCZ} and PRS_{EDUC} for the different DSM diagnosis groups.

EY, BACS and WRAT are positively related as expected; psychosis does not alter these correlations

An individual's educational attainment, cognitive functioning and intellectual potential are expected to be correlated with one another. The relationships among these three traits were similar in the two groups (Figure 2), indicating that the presence of psychosis did not alter the interdependence of these three phenotypes. Although BACS, WRAT and EY were significantly lower in the PSYCH group compared to the NPSYCH group, the effect size of deficit in BACS was three times as much as the deficits in EY or WRAT (BACS group difference Cohen's $d = -1.05$, $p = 2.3 \times 10^{-37}$, Figure S1). Additionally, partial correlation analyses between pairs of these three phenotypes controlling for the third phenotype revealed that, 1. EY and WRAT shared a positive correlation that could not be accounted for by BACS; 2. WRAT and BACS shared a positive correlation that could not be accounted for by EY; and 3. although EY and BACS were weakly positively correlated, this correlation was mediated via factors that could be captured by WRAT (Table S2).

Higher genetic risk of schizophrenia is associated with lower BACS score in nonpsychotic individuals

To evaluate whether the genetic risk of schizophrenia associates with variations in BACS, EY and WRAT, correlations of PRS_{SCZ} with these measures were calculated within the PSYCH and the NPSYCH groups separately. Since EY, BACS and WRAT were positively correlated (Figure 2), full correlations as well as partial correlations controlling for the remaining two traits were calculated. Figure 3 shows the strongest full and partial correlations for each phenotype. For all polygenic score correlation analyses the combined FDR corrected p-value threshold was determined to be $P_{\text{FDR-PRS}} = 0.0064$ (Methods). The numerical values for the correlation coefficients and the p-values for both groups at all P_T can be found in Table S3.

As shown in Figure 3, PRS_{SCZ} showed insignificant correlations with WRAT and EY in both the PSYCH and the NPSYCH groups. Significant negative correlations were observed between PRS_{SCZ} and BACS in the NPSYCH group but not in the PSYCH group at multiple P_T (Table S3). The strongest correlation, as shown in Figure 3a, was observed when EY and WRAT were regressed out at $P_T = 1 \times 10^{-4}$ ($r = -0.19$ and $p = 1.0 \times 10^{-4}$). Since the NPSYCH group was formed by combining the healthy control (HC) and the nonpsychotic family member (NPFAM) subgroups (Methods), HC/NPFAM status was accordingly used as a covariate in the above correlation analyses. Additionally, the relationship between PRS_{SCZ} and BACS was explored within the HC and the NPFAM subgroups separately (supplementary material) to ensure these results did not arise as an artifact of merging the two subgroups. There was no significant difference in the correlation coefficients in the two subgroups (Table S4). At the subgroup level, statistically significant correlation between BACS and PRS_{SCZ} was seen at $P_T = 10^{-4}$ in

the HC group ($r = -0.25$, $p = 1.9 \times 10^{-3}$), which remained significant when EY and WRAT were regressed out. In the NPFAM subgroup, significant negative correlation was detected at $P_T = 0.01$ ($r = -0.19$, $p = 6.4 \times 10^{-3}$) when EY and WRAT were regressed out (Table S4).

The PSYCH group consisted of schizophrenia, bipolar disorder and schizoaffective disorder probands. Correlation between BACS and PRS_{SCZ} within the schizophrenia probands only (SZP, $N = 100$) was also insignificant. Adding illness duration, number of hospitalization, chlorpromazine dose, number of psychotropic drugs, and social-functional scale score as covariates in the correlation analysis between BACS and PRS_{SCZ} did not alter the lack of significant results in the PSYCH group.

Higher genetic score of educational attainment is associated with higher EY and WRAT

To investigate the effects of the common genetic polygenic influences on educational attainment (36) on these traits, we analyzed the correlation of the polygenic score of educational attainment (PRS_{EDUC}) with EY, WRAT and BACS. As expected, PRS_{EDUC} showed significant positive correlations with EY in both the PSYCH group (Fig. 4b, strongest correlation of $r = 0.19$, $p = 0.0016$ at $P_T = 0.05$) and the NPSYCH group (Fig. 4a, strongest correlation of $r = 0.17$, $p = 7 \times 10^{-4}$ at $P_T = 0.01$). When BACS and WRAT were controlled for, these significant correlations decreased and the only remaining significant correlation was in the NPSYCH group at $P_T = 0.01$ ($r = 0.15$, $p = 0.003$). Statistically significant positive correlations were observed between PRS_{EDUC} and WRAT in both the PSYCH group and the NPSYCH group (Figs. 4a,b) at several thresholds

(strongest correlation of $r = 0.26$, $p = 10^{-5}$ at $P_T = 0.05$ in PSYCH and strongest correlation of $r = 0.15$, $p = 2.4 \times 10^{-3}$ at $P_T = 0.05$ in NPSYCH). No significant correlation was found between PRS_{EDUC} and BACS in either group. The numerical values for all the correlation coefficients and p-values can be found in Table S5. Together these results indicate that the common variants associated with educational attainment influence education and premorbid intellectual potential (as measured by WRAT) more significantly than general cognitive performance in the general population.

Discussion

Educational attainment (EY), premorbid intellectual potential (WRAT), and cognitive performance (BACS) are interrelated and correlated phenotypes (Figure 2), making it challenging and important to disentangle the effects of genetic risk on each. In the current work, in which we sought to disentangle these effects through simultaneous conditional analyses, the genetic risk of schizophrenia (PRS_{SCZ}) significantly associated only with BACS and not with EY or WRAT in nonpsychotic individuals. PRS_{SCZ} explained about 4% of the variance in the composite BACS scores in the NPSYCH group.

These results may help to understand earlier results on genetic correlations of schizophrenia with other phenotypes. Using the UK Biobank sample Hagenaars et al. (29) found significant negative relationships between schizophrenia genetic risk and three cognitive traits: verbal-numerical reasoning, memory and reaction time. While the composite BACS score is not directly comparable to these three phenotypes, the six subtests used to calculate the BACS composite score significantly overlap with these three cognitive subdomains (Methods). Hence, our findings are consistent with the results

in Hagenaars et al. We further found that the effect of the genetic risk of schizophrenia on BACS was significant after accounting for variability due to EY and WRAT. While BACS measures an individual's ability to use cognitive resources to solve problems, WRAT is more of a measure of crystallized verbal knowledge. These results indicate that cognitive domains measured by BACS— rather than other brain phenotypes that shape premorbid intelligence or educational attainment – are likely more direct targets of the genetic risk factors of schizophrenia

Though we observed a strong negative correlation between PRS_{SCZ} and BACS at multiple P_T in the nonpsychotic group (NPSYCH, Figure 3), as well as individually within the healthy controls and the nonpsychotic family members (Table S4), we did not observe such a correlation among psychosis patients (PSYCH). Thus, we found no evidence that psychotic individuals at elevated genetic risk (due to common variants) are more severely cognitively impaired than psychotic individuals at lower genetic risk. Cognitive deficits in the patient group may thus reflect morbid factors that are not predicted by PRS_{SCZ}. The morbid consequences of disease progression, protective effects of supportive care, and the effects of medications, medical and psychiatric comorbidity and substance use may have more profound effects on pathologic trajectories than any pre-existing variation in cognition, and may involve biological mechanisms distinct from those that mediated pre-onset risk. Recently, a similar result was reported in a study of Autism Spectrum Disorder (ASD) in which the polygenic risk of ASD did not predict IQ in the ASD probands (despite a strong positive correlation in the general population, 38) although the polygenic scores of educational attainment and schizophrenia did (39). This too could be due to non-genetic factors playing a significant role in determining cognitive

function in the ASD patients, overshadowing the cognitive variability due to the genetic risk of developing the disease.

A higher value of PRS_{EDUC} reflects the inheritance of a larger-than-average number of genetic factors (alleles) that associate with educational attainment in the broader population. Therefore, the observed significant positive correlations between PRS_{EDUC} and EY in both the PSYCH and the NPSYCH groups in our sample were expected (Figure 4). Additionally, we found that PRS_{EDUC} was positively associated with WRAT in both the PSYCH and NPSYCH groups. The correlation between PRS_{EDUC} and BACS was not significant in either group. Together, these results imply that the common variants associated with educational attainment influence EY and WRAT more significantly than they influence BACS in the general population. It should be noted that Hagenaars et al. (29) also reported lack of significant correlations between their educational attainment polygenic score and memory and reaction time. However they did report significant positive correlation of education polygenic score with verbal-numerical reasoning, which is likely closer to our measure of WRAT than to BACS. In our sample of 737 individuals, the absence of correlation between two quantities does not necessarily imply the true absence of effects. Our study might be underpowered to detect weaker correlations (at $P_{FDR}=0.0064$, 25% and 17% power to detect a correlation of 0.1 in the NPSYCH group and the PSYCH group, respectively).

Decreased performances on all three cognitive measures used in this work: BACS, WRAT and EY, could be considered as risk factors for future psychotic illness. However, the phenotypic deficit in BACS in psychotic probands is much more significant compared to the deficits in EY and WRAT. Additionally, while the cognitive domains measured by

BACS show increasing deficits along the psychosis onset trajectory during the high-risk prodromal phase, WRAT remains a stable measure during this period influenced primarily by events occurring in earlier life well before the illness prodrome. In this study, genetic risk of developing schizophrenia was more strongly connected to cognitive performance among unaffected individuals than to either (i) the magnitude of cognitive impairment among psychosis patients, or (ii) educational attainment or premorbid intellectual potential (as measured by WRAT reading scores). Our results indicate that BACS is genetically more closely associated with schizophrenia and possibly psychotic illnesses in general, compared to the other two measures in spite of the interrelated nature of these phenotypes. These results suggest that specific cognitive phenotypes, independent of education or general intelligence, could be more deeply studied for insight into the specific processes affected by the genetic influences on psychosis risk. Studies of the premorbid period and/or the high-risk prodromal stage may be particularly valuable for understanding the connections and sequences of events involving genetic risk, performance in specific cognitive domains, and chronic psychotic illness.

Methods

Study design and participants

This study includes 737 Caucasians (Table S6) from the Bipolar-Schizophrenia Network for Intermediate Phenotypes (B-SNIP), which is a multi-site consortium studying risk factors for psychosis (3, 37). Previous work using this cohort reported BACS performance to be consistent with a dimensional model of psychosis in which the schizophrenia probands were most severely impaired, followed by schizoaffective

disorder probands and bipolar disorder probands (3). Also, a recent study using the B-SNIP cohort reported the lack of a biological basis for the segmenting of the psychoses by DSM diagnoses (40). Due to these findings and also due to the relatively small sample size we combined all psychotic probands to form the PSYCH group (N = 314). To control for effects that might arise from differences among the schizophrenia/bipolar/schizoaffective disorder probands, the DSM diagnosis of all probands was used as a covariate for correlation analyses within the PSYCH group. The NPSYCH group consisted of nonpsychotic individuals and was created by combining healthy controls (HC, N = 180) and first-degree relatives with no history of psychosis (NPFAM, N = 243). It was shown in (3) that nonpsychotic relatives with elevated axis II traits (cluster A or cluster B) were significantly cognitively impaired compared to healthy controls. To minimize heterogeneity within the NPSYCH group, these individuals were excluded. Additionally, HC/NPFAM status as well as the respective proband's DSM diagnosis (for NPFAM samples only) were used as covariates for all correlation analyses within the NPSYCH group. In the families for which multiple family members' data were available, only one member selected at random from each family was included in the NPSYCH group.

Details of participant selection criteria can be found in previous publications from the B-SNIP consortium (3). Institutional review boards at each site approved the study and all sites used identical diagnostic, clinical, and recruitment techniques (37).

Cognitive measures

Cognitive function was measured by the Brief Assessment of Cognition in Schizophrenia (BACS), which is a 30-minute test of global neuropsychological function (31). The BACS consisted of six subtests: verbal memory (VM), digit sequencing (DS), verbal fluency (VF), token motor task (TM), digit symbol coding (DSC) and Tower of London (TL). These tests encompass four cognitive domains: verbal memory, processing speed, reasoning and problem solving, and working memory (3). Premorbid intellectual potential was measured using the reading score of the Wide Range Achievement Test (WRAT IV), which has a phenotypic correlation of ~ 0.4 with full-scale intelligent quotient (32, 41). Self-reported years of education completed at the time of recruitment was used as a measure of educational attainment (EY).

Collection and quality control of genetic data

Genetic data for the B-SNIP project were collected for 2053 subjects (multi-ethnic sample) using the Illumina Infinium PsychArray BeadChip™ platform. Genotypes underwent quality control using PLINK 1.9 (42, 43) based on a standardized protocol (44) in which individual markers were removed if they had a missing rate greater than 5%, deviated from Hardy-Weinberg equilibrium ($p < 10^{-6}$), had a very low minor allele frequency (< 0.01), or demonstrated significantly different call rates between psychiatric probands and controls ($p < 10^{-5}$). Subjects were removed for discordant sex information, outlying heterozygosity (> 3 standard deviations above the mean), or excessive missing genotype data (> 0.1). Kinship analyses were run using both PREST-plus (45) and KING (46) to confirm within-family kinships as well as to check for cryptic relatedness between unrelated individuals. Samples showing a relationship closer than 3rd degree to another

unrelated individual were excluded. Samples of relatives that failed or showed a misrepresented kinship, or monozygotic twins were also excluded. After these initial quality control steps 1962 multi-ethnic samples remained.

Ancestry Verification and Final Sample

Only samples with nonmissing age, sex, data collection site, BACS, WRAT, education, and history of psychosis were retained resulting in 1528 samples of whom 927 were self reported Caucasians (SRC). To avoid population stratification only SRC samples were used in all analyses. The ancestries of these SRC samples were verified by principal component analysis combining the B-SNIP genotype data with the 1000 Genomes phase 1 data (47). Samples that were more than four standard deviations away from the SRC group mean along the first ten principal components were excluded (N = 49). In the remaining sample only one first-degree relative of each proband was kept at random in order to avoid having related individuals in the NPSYCH group. Additionally, relatives with elevated cluster A or cluster B traits were removed. This final B-SNIP sample (N = 737) along with the different ethnic groups of the 1000 genome sample can be seen in the principal component scatter plot in Figure S4. The first ten principal components (PCs) of the ancestry analysis were used as covariates for all genetic correlation analyses to reduce the effects of genetic variability due to ancestry.

Imputation of genotyped data and polygenic score construction

Imputation of genetic data was performed using HAPI-UR for pre-phasing (48) and IMPUTE2 for imputation (49,50) using a multiethnic reference panel (51).

Chromosomes were phased separately, and then divided into 5-million-base-pair chunks for imputation. The 1000 Genomes phase 1 data were used as a reference panel for imputation (47). Poorly imputed single nucleotide polymorphisms (SNPs) were filtered post-imputation by removing SNPs that had information score less than 0.5 (52).

For a given set of SNPs, polygenic risk score was calculated for an individual as a weighted sum of that individual's number of risk alleles at each SNP in the SNP set, with each allele weighted by the log of the odds ratios from the discovery sample, using custom scripts. Schizophrenia polygenic profile scores (PRS_{SCZ}) and educational polygenic score (PRS_{EDUC}) were calculated using the summary statistics of the Psychiatric Genome Consortium (PGC) schizophrenia GWAS meta-analysis results (22) and the summary statistics from Okbay et al. (36), respectively. Both scores were calculated for seven p-value thresholds of significance of association: $P_T < 10^{-4}$, 0.001, 0.01, 0.05, 0.1, 0.5 and 1.0. Multiple values of P_T were investigated instead of a single threshold as the correlation analyses between the genetic profile scores and the cognition traits were exploratory. Scores calculated using higher P_T include more SNPs that are less significantly associated with the phenotype.

Statistical Analysis

All statistical analyses were performed using Matlab (2012b). Group differences were calculated after regressing out the effects of age, sex, data collection site and the first ten principal components from the genetic ancestry analysis. The group difference p-values were calculated using the non-parametric Kruskal-Wallis test and effect sizes were calculated using the standard pooled variance Cohen's d approach. Partial correlations

between two variables while controlling for other variables were calculated using the Spearman Rank method (“partialcorr” function in Matlab) to minimize unwanted effects of outliers. Age, sex, data collection site, the first 10 principal components from the genetic ancestry analysis, and DSM diagnosis (schizophrenia/bipolar disorder/schizoaffective disorder status for members of the PSYCH group and respective relative’s diagnosis for member’s of the NPFAM group) were regressed out for correlation analyses within each group. As an additional precaution, the samples’ HC/NPFAM status was used as a covariate for all analyses within the NPSYCH group.

To correct for multiple hypotheses testing a false discovery rate (FDR) approach was used (53) instead of the more severe Bonferroni correction (54) as many of the variables, including polygenic scores calculated at different P_T , were correlated. Analysis-specific FDR thresholds as well as a threshold combining all p-values in this work were calculated at level $\alpha = 0.05$. For the entire study $P_{FDR-ALL}$ was 0.011. For analyses with polygenic scores the combined $P_{FDR-PRS}$ was 0.0064 (Table S3, Table S4, Table S5). For group difference analyses in BACS, EY and WRAT between the PSYCH and the NPSYCH groups (Fig S1) and between the HC and the NPFAM groups (Fig S3) P_{FDR} was 0.022. P_{FDR} for group differences in PRS_{SCZ} and PRS_{EDUC} was 0.00089 (Table S1). For correlation analysis between EY, BACS, and WRAT (Table S2) P_{FDR} was 0.006. Significant results are reported based on analysis-specific FDR p-values since in most cases these values were smaller than $P_{FDR-ALL}$. Relevant significance thresholds are mentioned with each result.

Our sample size was not sufficiently large to implement the recently developed statistical genetics methods of Linkage Disequilibrium (LD) Score regression (55) and

Genome-Wide Complex Trait Analysis (GCTA, 56), so we focused on the polygenic score approach.

References

1. Kahn RS, Keefe RS (2013) Schizophrenia is a cognitive illness: time for a change in focus. *JAMA Psychiatry* 70(1):1107-1112.
2. Kraepelin E (1893) *Psychiatrie: Ein Lehrbuch für Studierende und Ärzte*. 4th ed. Leipzig, Germany: Verlag von Johann Ambrosius Barth; 1893:435-445.
3. Hill SK, et al. (2013) Neuropsychological impairments in schizophrenia and psychotic bipolar disorder: findings from the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). *Am J Psychiatry* 170(11):1275-1284.
4. Lewandowski KE, Cohen BM, Ongur D (2011) Evolution of neuropsychological dysfunction during the course of schizophrenia and bipolar disorder. *Psychol Med* 41(2):225-241.
5. Bilder RM, et al. (2000) Neuropsychology of first-episode schizophrenia: initial characterization and clinical correlates. *Am J Psychiatry* 157(4):549-559.
6. Reichenberg A, et al. (2010) Static and dynamic cognitive deficits in childhood preceding adult schizophrenia: a 30-year study. *Am J Psychiatry* 167(2):160-169.
7. Liu CH, Keshavan MS, Tronick E, Seidman LJ (2015) Perinatal risks and childhood premorbid indicators of later psychosis: next steps for early psychosocial interventions. *Schizophr Bull* 41(4):801-816.

8. Khandaker GM, Barnett JH, White IR, Jones PB (2011) A quantitative meta-analysis of population-based studies of premorbid intelligence and schizophrenia. *Schizophr Res* 132(2-3):220-227.
9. McCabe JH, et al. (2008) Scholastic achievement at age 16 and risk of schizophrenia and other psychoses: a national cohort study. *Psychol Med* 38(8):1133-1140.
10. Seidman LJ, et al. (2016) Association of neurocognition with transition to psychosis: baseline functioning in second phase of the North American Prodrome Longitudinal Study. *JAMA Psychiatry* 73(12):1239-1248.
11. Seidman LJ, et al. (2010) Neuropsychology of the prodrome to psychosis in the NAPLS consortium: relationship to family history and conversion to psychosis. *Arch Gen Psychiatry* 67(6):578-588.
12. Fusar-Poli P, et al. (2012) Cognitive functioning in prodromal psychosis: a meta-analysis. *Arch Gen Psychiatry* 69:562-571.
13. Simon AE, et al. (2007) Cognitive functioning in the schizophrenia prodrome. *Schizophr Bull* 33(3):761-777.
14. Niendam TA, et al. (2007) The course of neurocognition and social functioning in individuals at ultra high risk for psychosis. *Schizophr Bull* 33(3):772-781.
15. Sullivan PF, Kendler KS, Neale MC (2003) Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry* 60(12):1187-1192.
16. Bouchard TJ Jr (2014) Genes, evolution and intelligence. *Behav Genet* 44(6):549-577.
17. Haworth CM, et al. (2015) The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol Psychiatry* 15(11):1112-1120.

18. Deary IJ, Penke L, Johnson W (2010) The neuroscience of human intelligence differences. *Nat Rev Neurosci* 11(3):201-211.
19. Plomin R, Deary IJ (2015) Genetics and intelligence differences: five special findings. *Mol Psychiatry* 20(1):98-108.
20. Davies G, et al. (2015) Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=53949). *Mol Psychiatry* 20(2):183-192.
21. Davies G, et al. (2011) Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Mol Psychiatry* 16(10):996-1005.
22. Schizophrenia working group of the Psychiatric Genomics Consortium (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511(7510):421-427.
23. Psychiatric GWAS consortium bipolar disorder working group (2011) Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet* 43(10):977-983.
24. Touloupoulou T, et al. (2007) Substantial genetic overlap between neurocognition and schizophrenia: genetic modeling in twin samples. *Arch Gen Psychiatry* 64(12):1348-1355.
25. Touloupoulou T, et al. (2010) Impaired intellect and memory: a missing link between genetic risk and schizophrenia? *Arch Gen Psychiatry* 67(9):905-913.

26. Fowler T, Zammit S, Owen MJ, Rasmussen F (2012) 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* 69(5):460-466.
27. Bulik-Sullivan B, et al. (2015) An atlas of genetic correlations across human diseases and traits. *Nat Genet* 47(11):1236-41.
28. Hubbard L, et al. (2016) Evidence of common genetic overlap between schizophrenia and cognition. *Schizophr Bull* 42(3):832-842.
29. Hagenaars SP, et al. (2016) Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N=112151) and 24 GWAS consortia. *Mol Psychiatry* 21(11):1624-1632.
30. Lencz T, et al. (2014) Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: a report from the Cognitive Genomics consortium (COGENT). *Mol Psych* 19(2):168-174.
31. Keefe RS, et al. (2004) The Brief Assessment of Cognition in Schizophrenia: reliability, sensitivity and comparison with a standard neurocognitive battery. *Schizophr Res* 68(2-3):283-297.
32. Kareken DA, Gur RC, and Saykin AJ (1995) Reading on the Wide Range Achievement Test-Revised and parental education as predictors of IQ: comparison with the Barona formula. *Arch Clin Neuropsychol* 10(2):147-157.
33. Bright P, Jaldow E. and Kopelman MD (2002) The National Adult Reading Test as a measure of premorbid intelligence: a comparison with estimates derived from demographic variables. *J Int Neuropsychol Soc.* 8(6):847-854.
34. Gladsjo JA, Heaton RK, Palmer BW, Taylor MJ, Jeste DV (1999) Use of oral reading

to estimate premorbid intellectual and neuropsychological functioning. *J Int Neuropsychol Soc* 5(3):247-254.

35. Davies G, et al. (2016) Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112151). *Mol Psychiatry* 21(6):758-767.
36. Okbay A, et al. (2016) Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 533(7604):539-542.
37. Tamminga CA, et al. (2014) Bipolar and schizophrenia network for intermediate phenotypes: outcomes across the psychosis continuum. *Schizophr Bull* 40(6):S131-137.
38. Clarke TK, et al. (2016) Common polygenic risk for autism spectrum disorder (ASD) is associated with cognitive ability in the general population. *Mol Psychiatry* 21(3):419-425.
39. Weiner DJ, et al. (2016) Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Bioarxiv.org*, doi: <https://doi.org/10.1101/089342>.
40. Clementz BA, et al. (2016) Identification of distinct psychosis biotypes using brain-based biomarkers. *Am J Psychiatry* 173(4):373-384.
41. Germine L, et al. (2016) Association between polygenic risk for schizophrenia, neurocognition, and social cognition across development. *Transl Psychiatry* 6(10):e924 doi:10.1038/tp.2016.147.
42. Purcell S, et al. (2007) PLINK: a tool set for whole-genome association and population-based lineage analyses. *Am J Hum Genet* 81(3):559-575.

43. Chang CC, et al. (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4:7.
44. Anderson CA, et al. (2010) Data quality control in genetic case-control association studies. *Nat Protoc* 5(9):1564-1573.
45. Sun L, Dimitromanolakis A (2014) PREST-plus identifies pedigree errors and cryptic relatedness in the GAW18 sample using genome-wide SNP data. *BMC Proc* 8 (Suppl 1 Genetic Analysis Workshop 18 Vanessa Olmo): S23. doi: 10.1186/1753-6561-8-S1-S23. eCollection 2014.
46. Manichaikul A, et al. (2010) Robust relationship inference in genome-wide association studies. *Bioinformatics* 26(22):2867-2873.
47. 1000 Genomes Project Consortium, et al. (2012) An Integrated map of genetic variation from 1,092 human genomes. *Nature* 491(7422):56-65.
48. Williams AL, Patterson N, Glessner J, Hakonarson H, Reich D (2012) Phasing of many thousands of genotyped samples. *Am J Hum Genet* 91(2):238-251.
49. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR (2012) Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* 44(8):955-959.
50. Howie BN, Donnelly P, Marchini J (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *Plos Genet* 5(6):e1000529.
51. Howie B, Marchini J, Stephens M (2011) Genotype imputation with thousands of genomes. *G3 (Bethesda)* 1(6):457-470.

52. Marchini J, Howie B (2010) Genotype imputation for genome-wide association studies. *Nat Rev Genet* 11(7):499-511.
53. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 57(1):289-300.
54. Bonferroni, CE (1936) "Teoria statistica delle classi e calcolo delle probabilità." Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze 8:3-62.
55. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics Consortium., Patterson N, Daly MJ, Price AL, Neale BM (2015) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 47(3):291-295.
56. Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: A tool for genome-wide complex trait analysis. *Am J Hum Genet* 88(1):76-82.

Legends

Figure 1: Mean Polygenic scores of schizophrenia (PRS_{SCZ}) in the psychotic (PSYCH, $N = 314$) and the nonpsychotic (NPSYCH, $N = 423$) groups. The vertical black lines show the standard errors of the mean (SEM). Scores were calculated at seven p-value thresholds (P_T): 0.0001, 0.001, 0.01, 0.05, 0.1, 0.5 and 1.0, which are shown in different colors. All scores were z-transformed before mean and SEM calculation. PRS_{SCZ} was significantly higher ($p \leq P_{FDR} = 8.9 \times 10^{-4}$) in the PSYCH group compared to the NPSYCH group at all P_T . Table S1 shows the p-values for this analysis.

Figure 2: Relationship between BACS, educational attainment EY and WRAT. The three phenotypes were positively correlated in both the PSYCH ($N = 314$) and the NPSYCH ($N = 423$) groups, with the strongest correlation between BACS and WRAT ($r \sim 0.4$, $p < 10^{-11}$). The magnitudes of the correlations were not significantly different between the psychotic (PSYCH) and the nonpsychotic (NPSYCH) groups. More details can be found in Table S2.

Figure 3: Correlations of polygenic score of schizophrenia (PRS_{SCZ}) with BACS, WRAT and EY in the nonpsychotic (panel a, NPSYCH, $N = 423$) and the psychotic groups (panel b, PSYCH, $N = 314$). Only the strongest correlation for each phenotype is shown with the corresponding P_T labeled. Correlation coefficients are plotted with confidence intervals reflecting the FDR-corrected significance p-value of $P_{FDR-PRS} = 0.0064$. Confidence intervals not including the $r = 0$ line correspond to significant correlations. The effects of age, sex, collection site, ancestry principal components, DSM

diagnoses (in the PSYCH group) and healthy control (HC)/ nonpsychotic family member (NPFAM) status (within the NPSYCH group) were regressed out of EY, BACS, WRAT and PRS_{SCZ}. Within the NPFAM group, the respective proband's DSM diagnosis was also added as a covariate. Full correlations and partial correlations (in each case controlling for the other two traits) with PRS_{SCZ} are shown. Correlation coefficients and corresponding p-values for all P_T can be found in Table S3.

Figure 4: Correlations of polygenic score of educational attainment (PRS_{EDUC}) with BACS, WRAT and EY in the nonpsychotic (panel a, NPSYCH, N = 423) and the psychotic groups (panel b, PSYCH, N = 314). Only the strongest correlation for each phenotype is shown with the corresponding P_T labeled. Correlation coefficients are plotted with confidence intervals reflecting the FDR-corrected significance p-value of $P_{\text{FDR-PRS}} = 0.0064$. Confidence intervals not including the $r = 0$ line correspond to significant correlations. The effects of age, sex, collection site, ancestry principal components, DSM diagnoses (in the PSYCH group), and healthy control (HC)/ nonpsychotic family member (NPFAM) status (in the NPSYCH group) were regressed out of EY, BACS, WRAT and PRS_{SCZ}. Within the NPFAM group, the respective proband's DSM diagnosis was also added as a covariate. Full correlations and partial correlations (in each case controlling for the other two traits) with PRS_{EDUC} are shown. Correlation coefficients and corresponding p-values for all P_T can be found in Table S5.

Figure 1

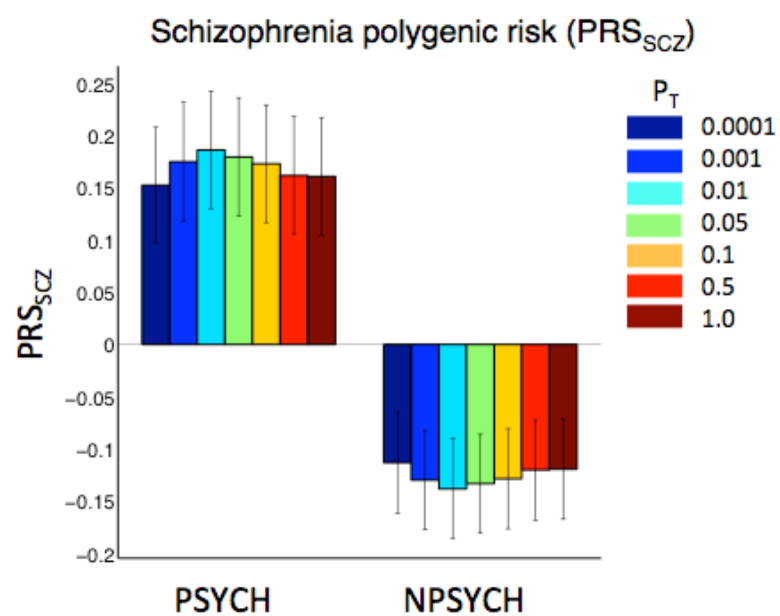


Figure 2

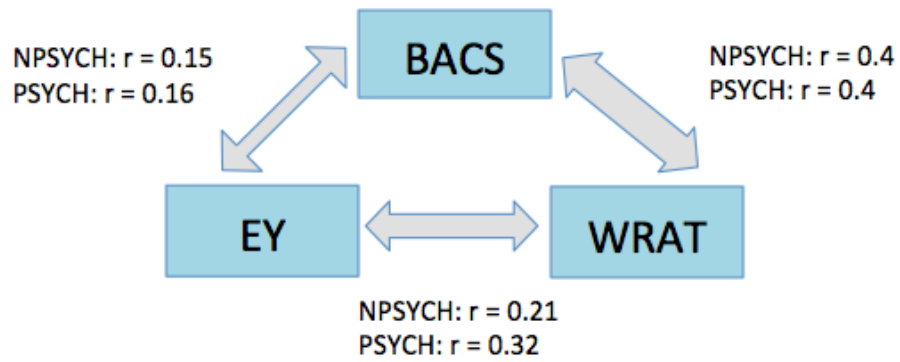


Figure 3

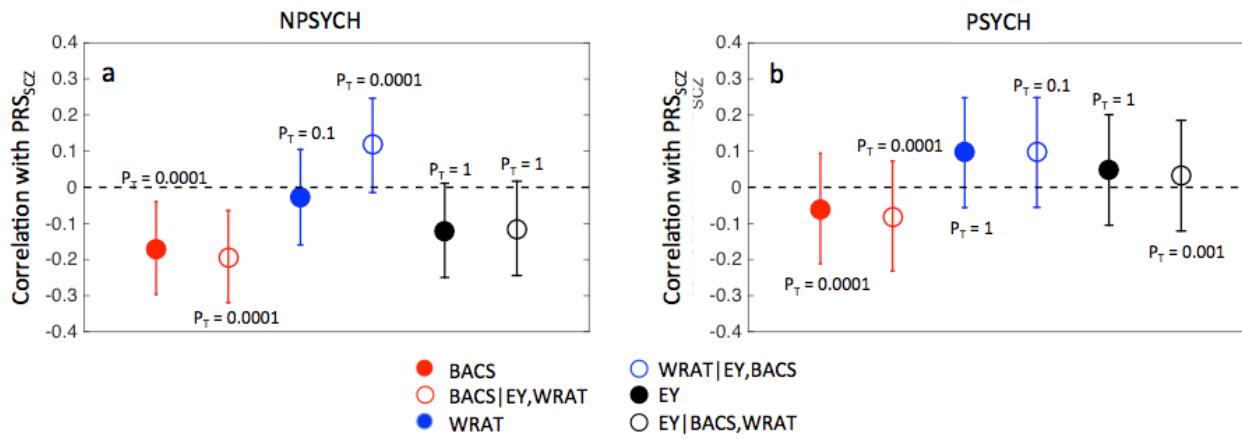


Figure 4

