Archival Report

Biological Psychiatry: GOS

Polygenic Scores Clarify the Relationship Between Mental Health and Gender Diversity

Taylor R. Thomas, Ashton J. Tener, Amy M. Pearlman, Katherine L. Imborek, Ji Seung Yang, John F. Strang, and Jacob J. Michaelson

ABSTRACT

BACKGROUND: Gender-diverse individuals are at increased risk for mental health problems, but it is unclear whether this is due to shared environmental or genetic factors.

METHODS: In two SPARK samples, we tested for associations of 16 polygenic scores (PGSs) with quantitative measures of gender diversity and mental health. In study 1, 639 independent adults (59% autistic) reported their mental health with the Adult Self-Report and their gender diversity with the Gender Self-Report (GSR). The GSR has 2 dimensions: binary (degree of identification with the gender opposite that implied by sex designated at birth) and nonbinary (degree of identification with a gender that is neither male nor female). In study 2 (N = 5165), we used a categorical measure of gender identity.

RESULTS: In study 1, neuropsychiatric PGSs were positively associated with Adult Self-Report scores: externalizing was positively associated with the attention-deficit/hyperactivity disorder PGS (β = 0.10 [0.03–0.17]), and internalizing was positively associated with the PGSs for depression (β = 0.07 [0–0.14]) and neuroticism (β = 0.10 [0.03–0.17]). Interestingly, GSR scores were not significantly associated with any neuropsychiatric PGS. However, GSR nonbinary was positively associated with the cognitive performance PGS (β = 0.11 [0.05–0.18]), with the effect size comparable in magnitude to the associations of the neuropsychiatric PGSs with the Adult Self-Report. Additionally, GSR binary was positively associated with the nonheterosexual sexual behavior PGS (β = 0.07 [0–0.14]). In study 2, the cognitive performance PGS effect replicated; transgender and nonbinary individuals had higher PGSs (t_{316} = 4.16).

CONCLUSIONS: We showed that while gender diversity is phenotypically positively associated with mental health problems, the strongest PGS associations with gender diversity were with the cognitive performance PGS, not the neuropsychiatric PGSs.

https://doi.org/10.1016/j.bpsgos.2024.100291

Sex and gender can have major impacts on health (1) (see Table 1 for our definitions). This stems from both extrinsic factors [e.g., health care barriers (2,3)] and biological factors, with sex and gender modulating the underlying molecular mechanisms of disease and well-being (4). In health research, sex is a more objective and well-defined variable than gender. This is because gender is often experienced on a continuum (5) and is multidimensional, with binary and nonbinary dimensions. Gender can be reported through self-endorsement of categorical gender identity labels, such as transgender, cisgender, nonbinary, and genderqueer. However, categorical labels may not be ideal for health research. Gender identity labels are contextually and culturally dependent (i.e., not accessible to all) and are often nonspecific in their meanings (6). Furthermore, gender diversity, which is a fundamental aspect of human diversity, is not only expressed by individuals with gender-diverse identities. People who identify as cisgender also exhibit some variation in dimensional gender diversity (7), but this diversity would be lost in studies that only use categorical gender identity labels. Therefore, parsing datasets based on numerous and nonspecific gender identity labels would erode the statistical power of research studies. A continuous, multidimensional characterization of gender that uses simple and widely accessible language will enable health researchers to appropriately incorporate gender diversity.

Gender diversity is a crucial variable to include in mental health research. Previous studies have reported higher rates of mental health problems in groups that express more gender diversity than the proportional cisgender majority, such as LGBTQ+ (lesbian, gay, bisexual, transgender, queer, +) individuals (9,10). One study found that LGBQ+ individuals had higher rates of anxiety, depression, and attempted suicide (11). LGBQ+ participants in the All of Us cohort (N = 329,038) had a higher prevalence of neuropsychiatric diagnoses (12). The exact mechanisms are not fully understood. However, research has shown that poorer mental health is at least partially due to factors related to experienced adversity from sexual orientation and/or gender diversity. Discrimination and resilience partially mediated negative mental health outcomes in LGBTQ+ college students (13). Access to gender-affirming

ISSN: 2667-1743

hormone therapy for transgender and gender nonbinary youth was associated with a reduced risk of depression and suicidality (14). To our knowledge, no study has used genetic data, so any possible genetic mechanisms are unknown.

Most behaviors are somewhat heritable, so we and others have hypothesized that gender diversity is also susceptible to genetic influences (15). A twin study of 4426 females estimated the heritability of adult gender expression (i.e., self-reported masculinity and femininity) at 11% and retrospective childhood gender typicality at 31% (16). However, searches for specific loci have been underpowered for gene discovery (17,18). Genome-wide association studies (GWASs) of human behavior often uncover many associated loci with small effects that contribute additively (19). Loci associated with 1 trait are often associated with others, which suggests that the 2 traits have a degree of pleiotropy. The predictive power of a GWAS polygenic score (PGS) depends on the power of the GWAS, which is driven chiefly by sample size. Among the well-powered GWASs, the most proximal trait to gender diversity is the nonheterosexual sexual behavior (NHSB) GWAS (20) that was performed with 408,995 UK Biobank participants. The trait was defined by participants' yes/no response to ever having sex with someone of the same sex (the nuance between same-sex and same-gender is lost due to the nature of the question). The estimated heritability of NHSB ranged from 8% to 25%. It was positively genetically correlated with several neuropsychiatric conditions and personality traits. However, interpretation of these genetic correlations is limited due to confounding with experienced adversity and psychiatric diagnoses.

In this study, we investigated whether gender diversity, like NHSB, is genetically associated with other behaviors and whether this plays a role in mental health. We invited a subset of SPARK (Simons Foundation Powering Autism Research) participants (21) to complete surveys on their mental health and gender identity. SPARK is a national genetic study of more than 300,000 participants with and without autism. Previous studies have shown that there is an enrichment of gender diversity in autism (22), and gender-diverse individuals have been found to have increased levels of clinically relevant autistic traits and an increased likelihood of an autism diagnosis (23). This makes SPARK a logical choice for investigating this topic. In our sample of 696 participants (n = 639 of European genetic ancestry), we calculated PGSs for 16 behavior traits. We also administered 2 psychometrically valid self-report tools. The first, the Adult Self-Report (ASR) (24), measures several mental health outcomes and adaptive behaviors. The second, the Gender Self-Report (GSR) (25), captures 2 quantitative dimensions of gender diversity: binary gender diversity, which is the extent to which a person experiences themselves as the other binary gender (i.e., different from their sex designated at birth), and nonbinary gender diversity, which is the extent to which a person experiences themselves as neither female nor male. Then we sought to answer the following questions: first, are ASR scores phenotypically associated with GSR scores? Second, are behavior-related PGSs associated with ASR and GSR scores? How are PGS associations different between the ASR and GSR? Lastly, do the PGS findings broadly replicate in a larger sample with a categorical gender identity phenotype? See Figure 1 for an overview of the study.

Table 1. Working Definitions of Terms

Term	Definition
Sex	Sex recorded around the time of birth based on physiological and anatomical sex characteristics; also referred to as designated sex, natal sex, assigned sex, or recorded sex; unless otherwise indicated, instances of "sex" in this work should be understood to mean sex at birth.
Gender Identity	An individual's own inner experience and personal sense of their gender—being a boy/man/male; girl/woman/female; or another gender (e.g., gender queer, gender fluid).
Transgender	A gender identity describing an individual whose gender is different from their sex at birth.
Cisgender	A gender identity describing an individual whose gender identity aligns with their sex at birth.
Nonbinary	An umbrella term encompassing those whose gender identity cannot be adequately described in a male-female axis; In some nomenclatures, this may include identities such as genderqueer, agender, gender fluid, third gender, and many others.
TGNB	A term to describe individuals whose gender differs from their sex at birth (i.e., not exclusively cisgender).
Gender Expression	The way an individual expresses aspects of their gender through physical appearance, clothing choice, accessories, and behavior.
Gender Diversity	An umbrella term used to describe divergence from gender identities, norms, and/or expressions often prescribed to those of the designated sex; this may be measured either in a categorical or a continuous manner.
Gender Dysphoria	Clinically relevant distress resulting from an incongruence between one's gender identity and designated sex at birth.
Sexual Orientation	The self-endorsed community labels(s) one finds representative of the gender(s) of their sexual and/or romantic attractions.

Some definitions are from (7) and (8).
TGNB, transgender and gender nonbinary.

METHODS AND MATERIALS

Sample Description

SPARK (21) is a nationwide autism study conducted in the United States with more than 300,000 participants. SPARK is approved by the Western Institutional Review Board (#20151664). For study 1, independent adults with or without autism were invited to participate in our SPARK Research Match. The Research Match was approved by the University of Iowa Institutional Review Board (#201611784). Those who consented to participate were asked to complete the GSR (25), ASR (24), and additional questions on sexual orientation, gender identity, and gender expression. The sample size was 818. We removed 9 individuals who had withdrawn from SPARK since the Research Match based on version 8 (n = 809). The final sample size was 696 after genetic data availability and quality control filtering. For study 2, we used the version 8 background history. Independent adult data were self-report, whereas child and sibling datasets were parent-report. We retained children who were 14 years or older whose cognitive impairment status at enrollment was not significantly below age.

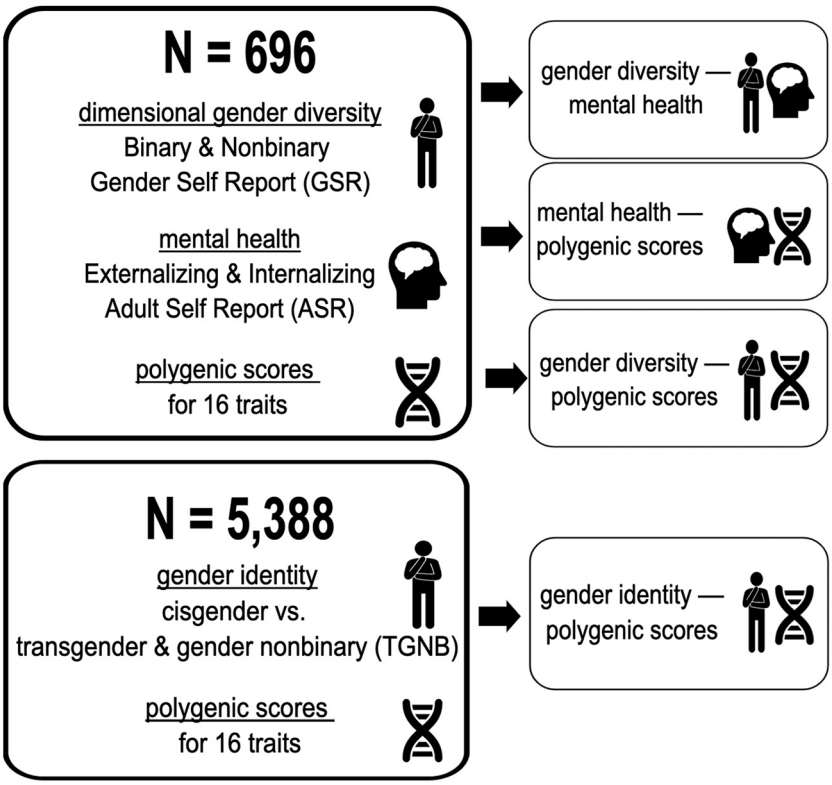

Figure 1. Overview of the study. In 696 participants (n = 639 of European genetic ancestry), dimensional gender diversity was measured using the Gender Self-Report (GSR), and mental health was measured using the Adult Self-Report (ASR). GSR and ASR scores were then tested for associations with 16 polygenic scores for psychiatric diagnoses, personality, and cognition. We then used categorical gender identity in 5388 participants (n = 5165 of European genetic ancestry) to test for replication of these polygenic score associations in the larger sample.

Study 1 Phenotypes

Labels of Gender Identity and Sexual Orientation. Participants were able to select as many labels for gender identity and sexual orientation as they found applicable. Selections of nonbinary, demigender, gender fluid, third gender, agender, gender neutral, pangender, bigender, and gender queer were categorized as nonbinary/neutral. Cisgender and transgender were each categorized separately. Participants who did not endorse any of the listed gender identities were excluded from analyses using gender identity labels (n = 66 of 696). For sexual orientation, participants who selected lesbian, gay, bisexual, pansexual, homosexual, queer, and/or polysexual were grouped as LGBQ+, and heterosexual orientation was classified separately. Participants who did not select any of the listed sexual orientation labels were excluded from analyses using sexual orientation labels (n = 73 of 696).

Gender Self-Report. The GSR item set was developed through an iterative multi-input, community-driven process with autistic cisgender, autistic gender-diverse, and nonautistic cisgender and gender-diverse collaborators (25); Open Science Framework Development Summary: https://osf.io/qh25d/?view_only=c0ce41d07bca4af1b792e074d51b7ded. The final GSR item set was composed of 30 questions. The GSR factor analysis and generation of binary and nonbinary factor scores are described in Strang *et al.* (25). In the genetic sample of N = 696, GSR scores were adjusted for age, sex designated at birth, and autism by linear regression residualization and then z scaled.

Adult Self-Report. The ASR (24) is a questionnaire consisting of 129 items that assesses a range of adaptive behaviors and mental health outcomes. From 809 participants, 5 were removed because they had 12 (approximately 10%) or more missing ASR items. In the remaining n = 804, 0.2% of the data were missing, which were imputed to the median. The 2 ASR subscales were externalizing (aggressive, rule-breaking, and intrusive behavior) and internalizing (anxious, withdrawn/depressed, and somatic complaints). In the genetic sample of N = 696, ASR scores were adjusted for age, sex, and autism by linear regression residualization and then z scaled.

Study 2 Phenotypes

If a participant's designated sex at birth (options: male or female) did not match their gender (options: male, female, or other), then the participant was classified as transgender or gender nonbinary (TGNB). Then we merged with our Research Match gender identity labels. The final sample size was 5388, with 590 from the Research Match and the other 4798 from the background history.

Genotype Quality Control and Imputation

We used the genotype arrays from SPARK integrated whole-exome sequencing (iWES1) 2022 release and SPARK whole-genome sequencing (WGS) releases 2, 3, and 4. iWES1 (n = 69,592) was quality controlled on release, including removing samples due to heterozygosity or high missingness, so no quality

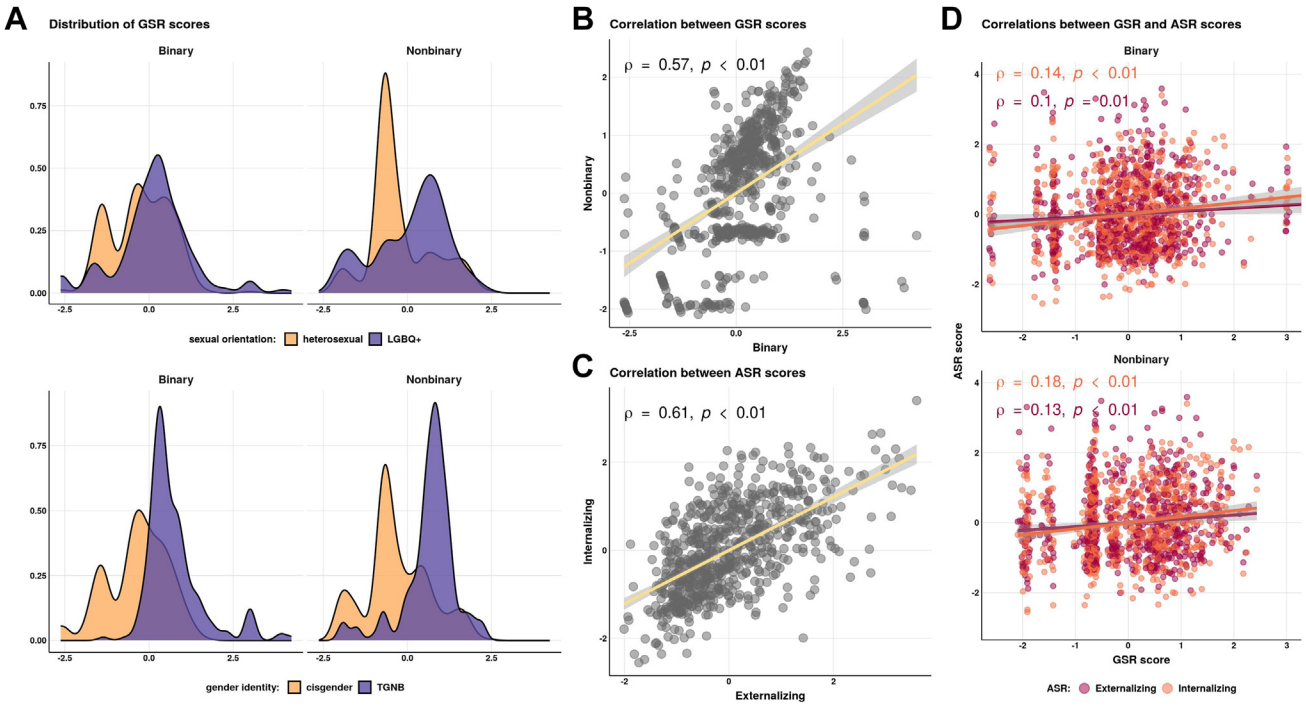

Figure 2. Distributions and correlations of Gender Self-Report (GSR) and Adult Self-Report (ASR) scores. **(A)** Distribution of the 2-dimensional gender diversity measures from the GSR: binary and nonbinary gender diversity. The sample size was $N = 696$. GSR scores were adjusted for age, sex, and autism. Histograms are colored by self-reported sexual orientation labels (top panel) and gender identity labels (bottom panel). Distributions of the GSR scores before adjusting for age, sex, and autism are shown in Figure S1. **(B)** Correlation of GSR scores. **(C)** Correlation of mental health measures from the ASR: externalizing and internalizing. ASR scores were also adjusted for age, sex, and autism. **(D)** Correlations between GSR scores and ASR scores. TGNB, transgender and gender nonbinary.

control was performed by us before imputation. iWES1 provided genetic ancestry assignments based on 1000 Genomes populations (26). WGS release 2 ($n = 2365$), release 3 ($n = 2871$), and release 4 ($n = 3684$) were not quality controlled on release, so we performed quality control using PLINK (27) before imputation. First, we removed participants from WGS if they were in iWES1. Second, we removed variants with missingness >0.1 and participants with missingness > 0.2. Third, we merged the 3 releases and removed any participant whose heterozygosity (F statistic) was not within 3 SDs of the mean. We used the TopMed reference panel (28) to identify strand flips. The final sample size for WGS 2 to 4 was $n = 8152$. iWES1 and WGS 2 to 4 were then imputed to TopMed (28) using the Michigan Imputation Server (29) with phasing and quality control steps included and to output variants with imputation quality $r^2 > 0.3$. After imputation, variants were filtered to HapMap single nucleotide polymorphisms ($n = 1{,}054{,}330$ variants) with imputation quality $r^2 > 0.8$ using bcftools (30). They were lifted over from hg38 to hg19 using the VCF-liftover tool (https://github.com/hmgu-itg/VCF-liftover) and normalized to the hg19 reference genome. Finally, files were merged, and variants with 0% missingness were retained ($n = 914{,}328$).

Genetic Ancestry

Genetic principal components (PCs) were calculated using bigsnpr (31), specifically following the authors' recommendations (32) and tutorial: https://privefl.github.io/bigsnpr/articles/bedpca.html. In summary, we 1) used the snp_plinkKINGQC function to identify and remove related participants at KING threshold of $2^{-3.5}$, 2) performed PC analysis using the bed_autoSVD with unrelated participants, 3) detected and removed PC outliers, 4) recalculated PCs, and 5) projected PCs onto the entire cohort using the bed_projectSelfPCA function. We used 40 PCs and performed k-means clustering with $K = 5$ [for the 5 populations of 1000 Genomes (26)] and used the genetic ancestry labels from iWES1 to assign labels to the genetic population clusters.

Relatedness

For study 1, from the 809 Research Match participants who completed the GSR, 804 completed the ASR and 727 had genetic data. This subset was pruned to remove related participants using GCTA (33) with a relatedness threshold of 0.125, corresponding to approximately third-degree relatives ($n = 31$ removed). For study 2, we retained only 1 participant from each family, with prioritization toward TGNB identities, and then removed related participants at the same threshold.

Polygenic Scores

PGSs were calculated using LDpred2 (34) and the bigsnpr tools (31) in R. Because SPARK is family based, an external linkage disequilibrium reference based on 362,320 participants in the UK Biobank (provided by the authors of LDpred2) was

used to calculate infinitesimal beta weights. PGSs were calculated from the following GWASs: attention-deficit/hyperactivity disorder (35), anorexia nervosa (36), autism (37), bipolar disorder (38), major depression (39), obsessive-compulsive disorder (40), schizophrenia (41), cognitive performance (42), educational attainment (42), and NHSB (20). The public LDpred2 beta weights from the Polygenic Index Repository (43) were used to calculate PGSs for extraversion (44), neuroticism (45), openness (46), risky behavior (47), number of children ever born (men) (48), and number of children ever born (women) (48). We residualized 20 genetic PCs from PGSs. We also accounted for age, sex, and autism using linear regression residualization. Lastly, PGSs were *z* scaled. PGS processing was done separately for study 1 and study 2.

We tested for PGS associations with GSR and ASR scores using linear models with covariates, as well as partial Spearman correlations (residualizing the covariates from PGSs and phenotypes prior to the correlation tests). The covariates were age, sex, and autism (linear model: phenotype ~ PGS + age + sex + autism), and the covariates were *z* scaled prior to model input. The pwr.r.test() (49) was used to determine the statistical power of the correlations. *p* Values were adjusted for multiple tests of 16 PGSs ($p_{adj.}$) using the Bonferroni method. We tested for PGS-by-GSR interactions with 3 linear models and then performed stratified correlations. The first model included the covariates as both main effects and interactions with the PGSs and GSR scores, as recommended by Keller (50) (ASR ~ PGS + GSR + PGS × GSR + age + sex + autism + PGS × age + PGS × sex + PGS × autism + GSR × age + GSR × sex + GSR × autism). For the second model, we wanted to be consistent with the previous analyses, so we tested interactions with the variables adjusted for covariates before model input (ASR ~ PGS + GSR + PGS × ASR). For the third model, we binarized the PGSs into the upper quartile (coded as 1) and the lower quartile groups (coded as 0), with $n = 160$ in each group and the middle 50% removed. This model was specified as ASR ~ PGSgroup + GSR + PGSgroup × GSR, again with the ASR and GSR scores adjusted for covariates prior to model input. To further investigate the interactions from the third model, we ran ASR-GSR correlations stratified by PGS groups.

RESULTS

Phenotypic Associations Between Dimensional Gender Diversity and Mental Health

The demographic characteristics of the SPARK Research Match participants are presented in Table 2. The sample size was 696 ($n = 639$ European genetic ancestry). Approximately one-third identified as TGNB. Fifty-eight percent were autistic, and 22% were male.

The 2 gender diversity scores, binary and nonbinary gender diversity, were from the GSR factor analysis (25). GSR scores range from 0 (no gender diversity) to 1 (high gender diversity) (Figure S1). These factor scores were adjusted for age, sex designated at birth, and autism by linear regression residualization and then *z* scaled. The distributions of these adjusted scores were affected by sexual orientation and gender identity, with higher scores in LGBQ+ and TGNB participants (Figure 2A). GSR binary and GSR nonbinary were positively correlated, $\rho = 0.57$, $p < .001$ (Figure 2B).

Table 2. Demographic Characteristics of Participants

	Study 1		Study 2	
Variable	*n*	%	*n*	%
Total sample size	696	–	5388	–
Age, Years, Mean	37	–	25	–
Sex, Male	153	22%	3271	61%
Autistic	403	58%	4731	88%
Gender Identity				
Cisgender identity	424	61%	5089	94%
TGNB identity	206	30%	299	6%
No gender identity label(s)	66	9%	–	–
Overlap with Research Match	–	–	590	11%
Genetic Population				
Africa	0	0%	0	0%
Americas	27	4%	100	2%
East Asia	4	1%	4	0%
Europe	639	92%	5165	96%
South Asia	26	4%	119	2%

Study 1 used 2 continuous measures of gender diversity from the Gender Self-Report as the phenotypes, and study 2 used a categorical gender identity phenotype.

TGNB, transgender and gender nonbinary.

The 2 mental health scores, externalizing and internalizing, are from the ASR (24). The ASR scores were also adjusted for age, sex, and autism prior to running the phenotypic correlations. ASR externalizing and ASR internalizing were positively correlated, $\rho = 0.61$, $p < .001$ (Figure 2C). ASR scores were positively correlated with GSR scores (Figure 2D). GSR binary was more strongly correlated with ASR internalizing ($\rho = 0.14$, $p < .001$) than with ASR externalizing ($\rho = 0.10$, $p = .01$). GSR nonbinary was also more strongly correlated with ASR internalizing ($\rho = 0.18$, $p < .001$) than with ASR externalizing ($\rho = 0.13$, $p < .001$).

PGS Associations With Dimensional Gender Diversity and Mental Health

We tested for associations between GSR and ASR scores and 16 behavior-related PGSs (Figure 3). PGS β effects are reported with 95% CIs from linear models with age, sex, and autism as covariates, with all variables *z* scaled. Partial correlations (residualizing the covariates from the phenotypes and PGSs) are shown in Table S1. Tests were performed in the European subset ($n = 639$), which has 80% power at $\alpha = .05$ to detect effects $\rho > \pm 0.11$, meaning that an absence of significant effects must be interpreted carefully. *p* Values were adjusted for multiple tests of 16 PGSs ($p_{adj.}$) using the Bonferroni method.

ASR scores had unsurprising positive associations with neuropsychiatric PGSs. ASR externalizing was positively associated with the attention-deficit/hyperactivity disorder PGS (β = 0.11 [0.05–0.18], $p = .008$, $p_{adj.} = .13$). ASR internalizing was positively associated with the depression PGS (β = 0.07 [0–0.14], $p = .041$, $p_{adj.} = .66$) and the neuroticism PGS (β = 0.10 [0.03–0.17], $p = .006$, $p_{adj.} = .10$).

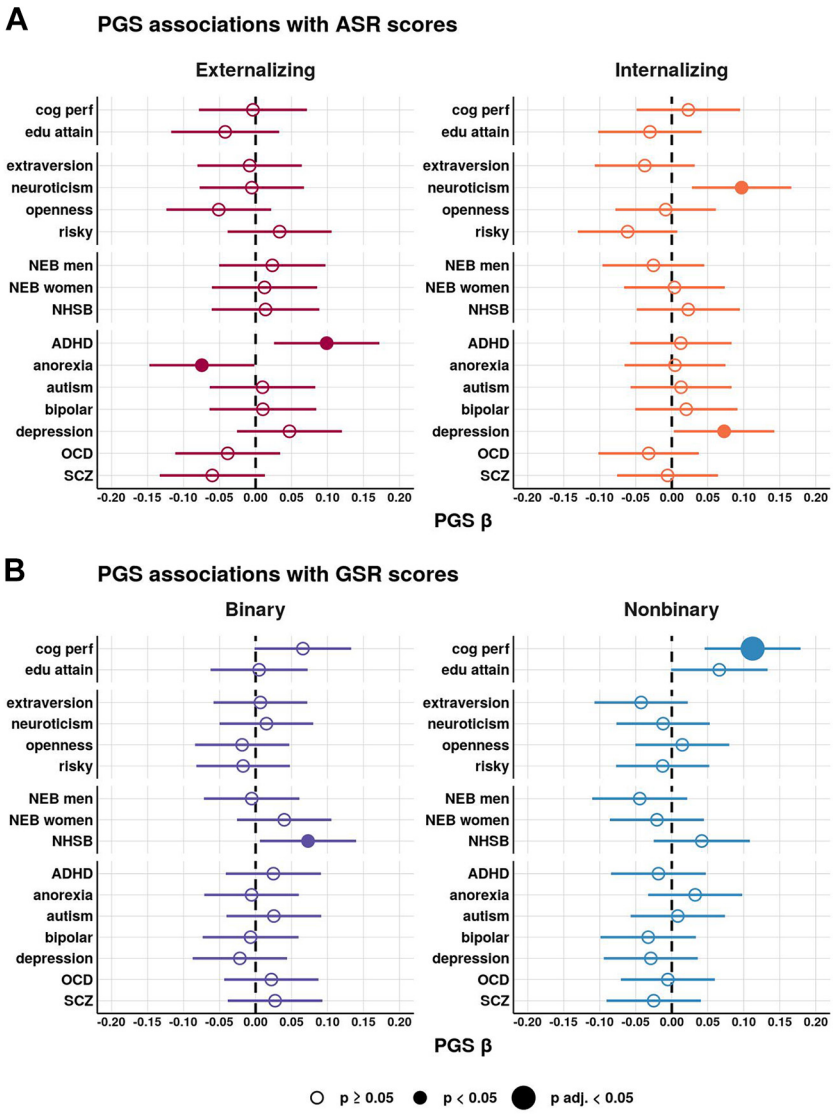

Figure 3. Polygenic score (PGS) associations with Gender Self-Report (GSR) and Adult Self-Report (ASR) scores. PGS associations with **(A)** ASR scores and **(B)** GSR scores. PGS β effects are shown from the linear model with age, sex, and autism as covariates (phenotype ~ PGS + age + sex + autism). PGS associations are from the European subset (n = 639). See Figure S2 for PGS associations in the entire sample (N = 696), and see Figure S3 for PGS associations with the sample stratified by autism diagnosis. Partial correlations (residualizing the covariates from the phenotypes and PGSs) are shown in Table S1. Nominal p values were adjusted for multiple tests of 16 PGSs using the Bonferroni method. ADHD, attention-deficit/hyperactivity disorder; NEB, number of children ever born; NHSB, nonheterosexual sexual behavior; OCD, obsessive-compulsive disorder; risky, risky behavior; SCZ, schizophrenia.

As expected, the NHSB PGS was positively associated with GSR binary (β = 0.07 [0.01–0.14], p = .033, $p_{adj.}$ = .53). Strikingly, the cognitive performance PGS was significantly positively associated with GSR nonbinary (β = 0.11 [0.05–0.18], $p <$ .001, $p_{adj.}$ = .02), meaning that polygenic propensity for higher cognitive performance was associated with elevated gender diversity. This PGS effect size was similar in magnitude to the associations of the neuropsychiatric PGSs with ASR scores. The cognitive performance PGS was also positively associated with GSR binary, but to a lesser extent that did not reach nominal significance (β = 0.07 [0–0.13], p = .06, $p_{adj.}$ = .89). No psychiatric PGSs were significantly associated with GSR scores.

Results were comparable when not filtering by genetic ancestry (N = 696, Figure S2). The cognitive performance PGS was positively associated with GSR nonbinary (β = 0.11 [0.04–0.17], $p <$ .001, $p_{adj.}$ = .01), and NHSB was positively associated with GSR binary (β = 0.07 [0–0.13], p = .04, $p_{adj.}$ = .61).

We tested whether correlations trended in the same direction when run separately in the autistic subset (n = 376) and among those without an autism diagnosis (n = 263) (Figure S3). These tests were not well powered, but the cognitive performance PGS had a trend-level positive association with GSR nonbinary in the autistic subset (β = 0.13 [0.04–0.22], p = .003, $p_{adj.}$ = .052) and in the nonautistic subset (β = 0.08 [−0.03 to 0.18], p = .14, $p_{adj.}$ = 1).

Replication of PGS Associations With a Categorical Gender Identity Phenotype

Next, we tested whether PGS associations were similar in a larger sample (N = 5388, n = 5165 European genetic ancestry) using a categorical gender identity phenotype (study 2). We used the background history to label individuals as cisgender or TGNB by discordance between the participant's designated sex at birth and their gender (options: male, female, or other).

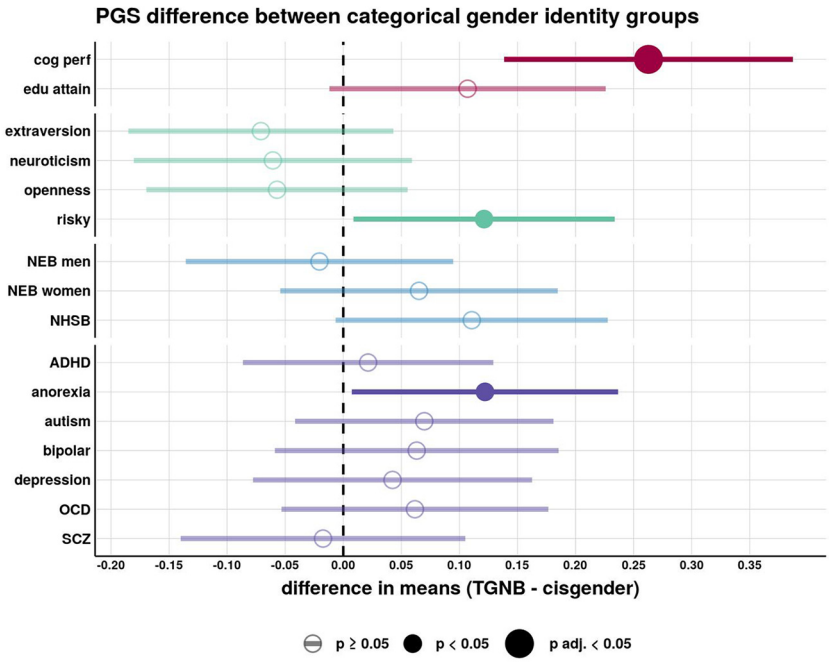

Figure 4. Replication of polygenic score (PGS) associations using a categorical gender identity phenotype in the larger sample (study 2). PGS difference in means (*t* tests) between gender identity groups in the larger cohort of European genetic ancestry (n = 5165). The 2 gender identity groups are cisgender (n = 4879) vs. transgender and gender nonbinary (TGNB, n = 286). PGSs were adjusted for age, sex, and autism prior to performing the *t* tests. See Figure S4 for PGS associations in the entire sample (N = 5388), and see Figure S5 for associations using PGSs that were not adjusted for age, sex, or autism. Nominal p values were adjusted for multiple tests of 16 PGSs using the Bonferroni method. ADHD, attention-deficit/hyperactivity disorder; NEB, number of children ever born; NHSB, non-heterosexual sexual behavior; OCD, obsessive-compulsive disorder; risky, risky behavior; SCZ, schizophrenia.

The mean age was 25 years, and 88% of participants were autistic.

We tested for PGS differences between the 2 gender identity groups with *t* tests (Figure 4). The strongest effect was observed for the cognitive performance PGS, with the TGNB group being significantly higher (x_d = 0.26 [0.14–0.39], t_{316} = 4.16, $p < .001$, $p_{adj.} < .001$). The TGNB group also had a significantly higher PGS for risky behavior (x_d = 0.12 [0.01–0.23], t_{325} = 2.12, p = .04, $p_{adj.}$ = .56) and anorexia (x_d = 0.12 [0.01–0.24], t_{323} = 2.09, p = .04, $p_{adj.}$ = .59). The NHSB PGS was close to nominal significance (x_d = 0.11 [−0.01 to 0.23], t_{321} = 1.86, p = .06, $p_{adj.}$ = 1).

We repeated tests without filtering by genetic ancestry (N = 5388) (Figure S4). The strongest effect was still for the cognitive performance PGS, which was higher in the TGNB group (x_d = 0.24 [0.12–0.36], t_{331} = 3.86, $p < .001$, $p_{adj.}$ = .002). We repeated the analysis without adjusting the PGS for covariates (Figure S5).

Interactions Between Dimensional Gender Diversity, Mental Health, and PGSs

Although our sample size was not well powered to detect interactions, having found little evidence for main effect PGS associations that explained the mental health associations with gender diversity, we decided to investigate interactions between PGSs and gender diversity. We tested PGS-by-GSR interactions with 3 linear models and performed correlations stratified by PGS (Figure S6). The first model included the covariates (age, sex, and autism) as main effects and interactions with PGS and GSR, as recommended by Keller (50). For the second model, we maintained consistency with the previous analysis that adjusted for covariates prior to testing. For the third model, we binarized PGSs into upper quartile (coded as 1) and lower quartile groups (coded as 0), with n = 160 in each group and the middle 50% removed. From the third model, we identified 3 nominally significant PGS-by-GSR interactions, specifically the schizophrenia and depression PGSs, although these interaction effects were not significant after adjusting for multiple testing. The most prominent interaction was the nonbinary and the schizophrenia PGS interaction on internalizing (β = 0.32 [0.1–0.53], $p < .001$, $p_{adj.}$ = .06). Within the entire sample of n = 639, GSR nonbinary and ASR internalizing were positively correlated, ρ = 0.18, $p < .001$. However, this apparent main effect appears to be driven by a context-specific interaction with the PGS: in the subset of participants who were highest on the schizophrenia PGS (e.g., upper quartile, n = 160), the correlation between GSR nonbinary and ASR internalizing was ρ = 0.36, $p < .001$, while in the lowest risk group (e.g., lower quartile, n = 160), there was no correlation, ρ = 0.02, p = .81. The interaction between the depression PGS and GSR nonbinary on ASR internalizing was also similar (β = 0.23 [0.02–0.45], p = .04, $p_{adj.}$ = .56).

DISCUSSION

Our analyses are the first to address relationships of multidimensional gender diversity with mental health and genetics. We used 2 quantitative measures of gender diversity, binary and nonbinary gender diversity from the GSR, in a neurodiverse sample of 696 adults in SPARK (21). We found greater gender diversity in female, autistic, and LGBTQ+ participants. Due to the structure of SPARK recruitment, we were only able to collect data from independent adults with autism or nonautistic immediate family members of someone with autism (mainly parents). Therefore, the elevated gender diversity in the autistic participants should be interpreted with the caveat that the nonautistic participants were older and were presumed to

adhere to more traditional gender roles. However, these results are consistent with previous research showing enrichment of gender diversity in autism (22). Intriguingly, while our results showed higher gender diversity in LGBTQ+ participants, many people who were cisgender also showed evidence of gender diversity, although not enough to report TGNB identities. This underscores the value of the GSR for capturing dimensional gender diversity beyond self-endorsed identities alone.

We tested 16 behavior-related PGSs for association with 2 GSR dimensions, and strikingly, the strongest association was the positive association between the cognitive performance PGS and both GSR binary and GSR nonbinary (Figure 3A). This finding was validated in our larger sample of n = 5165 with a categorical gender identity phenotype; the cognitive performance PGS was higher in the TGNB group than in the cisgender group (Figure 4). This suggests that cognitive capacity may be an important component in the development of more complex and nuanced gender identities. Beyond cognitive performance, the NHSB PGS was positively correlated with GSR binary. Although gender identity and sexual orientation are distinct, the NHSB GWAS is a well-powered GWAS that is adjacent to gender diversity. Recent research has found that just within heterosexuals, an NHSB PGS was positively associated with an increased number of partners (51). Building on this, our results suggest that gender diversity may be part of a pleiotropic ensemble of traits with adaptive advantages (e.g., cognitive performance).

We also expected neuropsychiatric PGSs to be positively associated with the GSR given that NHSB is positively genetically correlated with several neuropsychiatric conditions (20). Our sample size was on the low end for PGS associations: n = 639 provides 80% power for detecting $\rho \pm 0.11$, meaning that we are not powered to detect small effects. However, considering this previous research, it was surprising that we found no strong, significant positive associations of the neuropsychiatric PGSs with scores. This suggests that, within the statistical power limits of our sample, gender diversity may not have a strong direct genetic relationship with adult-onset psychiatric disorders. Instead, in our sample, higher gender diversity had the strongest genetic relationships with higher cognitive ability and NHSB.

The lack of a main genetic effect that links psychiatric conditions and gender diversity combined with our observation that GSR scores showed significant correlations with poorer self-reported mental health prompted us to examine the possibility of a relationship between gender diversity and mental health that depends on the level of genetic risk (i.e., interaction between a PGS and gender diversity). We observed differences in correlations when stratifying the sample by the schizophrenia and depression PGSs (Figure S6D). Groups with high depression and schizophrenia PGSs had the strongest GSR-ASR correlations, whereas GSR-ASR correlations in low PGS groups were absent (i.e., not nominally significant). This suggests that the PGSs for depression and schizophrenia may interact with gender diversity (or related environmental factors such as discrimination and/or minority stress), ultimately affecting mental health. In other words, the observed relationship between gender diversity and mental health may not be solely environmental or genetic but rather an interaction of the two. However, our sample size limits our ability to detect PGS main effects, let alone interactions, and therefore interaction effects must be interpreted with the understanding that they are small effects and are not significant after multiple testing correction.

Our results and interpretations have several limitations. The primary limitation is the small sample size, and we were only powered to detect strong PGS effects. In addition, age, sex, and autism were entangled with other variables of interest. Autism is confounded at the genetic level, as has been observed in other studies that have shown that educational attainment (37) and cognitive performance (52) are positively genetically correlated with autism. However, we repeated our analyses without adjusting for the PGS and phenotypes for autism (Table S1 and Figure S5) and also stratifying by autism (Figure S3) and found the results to be robust against inclusion or omission of autism.

Conclusions

In summary, our findings show that gender diversity as captured by the GSR had dimensional properties that share common genetic factors with cognitive performance and NHSB. Consistent with previous studies, we found that higher gender diversity was correlated with poorer mental health, but our results suggested that any polygenic contribution of psychiatric risk alleles to gender diversity, if such contributions exist, are not large. Instead, a person's polygenic background may function as a risk/resilience mechanism that interacts with gender diversity (and/or the adversity that comes with it) in shaping mental health outcomes.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by National Institutes of Health and National Institute of Mental Health (Grant No. R01HG012697 [to JJM and JFS], Grant No. MH105527 [to JJM], Grant No. DC014489 [to JJM], and Grant No. R01MH100028 [to JFS]), as well as Simons Foundation (Grant No. SFARI 516716 [to JJM]), Clinical and Translational Science Award (No. KL2TR001877 [to JJM]), the Fahs-Beck Fellow Grant (to JFS), and National Institutes of Health predoctoral training grant (Grant No. T32GM008629 [to TRT]). The Roy J. Carver Charitable Trust supports the work of JJM. This work was also supported by the University of Iowa Hawkeye Intellectual and Developmental Disabilities Research Center through Eunice Kennedy Shriver National Institute of Child Health and Human Development (Grant No. P50HD103556).

The study was designed by TRT, JFS, and JJM. The GSR scores were generated by JSY and JFS. The PGSs were generated by TRT and JJM. The analyses were performed by TRT, AJT, and JJM. The manuscript writing was done by all authors.

We thank our community advisory council, including members Elizabeth Graham, Sascha Klomp, and Jillian Nelson for their feedback throughout the research and writing process. We are also grateful to all the participants and families in SPARK, SPARK clinical sites, and SPARK staff. We appreciate obtaining access to genetic and phenotypic data for SPARK data on SFARI Base.

A previous version of this article was published on medRxiv: https://doi.org/10.1101/2021.11.22.21266696.

SPARK genetic data can be obtained at SFARI Base: https://base.sfari.org. SPARK Research Match data will be available to qualified approved researchers through SFARI Base after publication of this article. The code for all analyses can be found at https://research-git.uiowa.edu/michaelson-lab-public/gsr-polygenic-scores.

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Department of Psychiatry, University of Iowa, Iowa City, Iowa (TRT, AJT, JJM); Department of Urology, University of Iowa, Iowa City, Iowa

(AMP); Department of Family Medicine, University of Iowa, Iowa City, Iowa (KLI); Department of Human Development and Quantitative Methodology, University of Maryland, College Park, Maryland (JSY); Gender and Autism Program, Center for Neuroscience, Children's National Hospital, George Washington University School of Medicine, Washington, District of Columbia (JFS); Iowa Neuroscience Institute, University of Iowa, Iowa City, Iowa (JJM); and Hawkeye Intellectual and Developmental Disabilities Research Center, University of Iowa, Iowa City, Iowa (JJM).

Address correspondence to Jacob J. Michaelson, Ph.D., at jacob-michaelson@uiowa.edu.

Received May 10, 2023; revised Dec 21, 2023; accepted Dec 26, 2023.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsgos.2024.100291.

REFERENCES

1. Westergaard D, Moseley P, Sørup FKH, Baldi P, Brunak S (2019): Population-wide analysis of differences in disease progression patterns in men and women. Nat Commun 10:666.
2. Osika Friberg IO, Krantz G, Määttä S, Järbrink K (2016): Sex differences in health care consumption in Sweden: A register-based cross-sectional study. Scand J Public Health 44:264–273.
3. Romanelli M, Hudson KD (2017): Individual and systemic barriers to health care: Perspectives of lesbian, gay, bisexual, and transgender adults. Am J Orthopsychiatry 87:714–728.
4. Khramtsova EA, Davis LK, Stranger BE (2019): The role of sex in the genomics of human complex traits. Nat Rev Genet 20:173–190.
5. Lloyd AE, Galupo MP (2019): What people with normative identities believe about sex, gender and sexual orientation. Psychol Sex 10:269–280.
6. Morgan RE, Dragon C, Daus G, Holzberg J, Kaplan R, Menne H, *et al.*: Updates on terminology of sexual orientation and gender identity survey measures. FCSM 20–03. Federal Committee on Statistical Methodology. Available at: https://nces.ed.gov/FCSM/pdf/FCSM_SOGI_Terminology_FY20_Report_FINAL.pdf. Accessed November 1, 2021.
7. Thomas TR, Hofammann D, McKenna BG, I R van der Miesen A, Stokes MA, Daniolos P, Michaelson JJ (2020): Community attitudes on genetic research of gender identity, sexual orientation, and mental health. PLoS One 15:e0235608.
8. American Psychiatric Association: A guide for working with transgender and gender nonconforming patients. Available at: https://www.psychiatry.org/psychiatrists/cultural-competency/education/transgender-and-gender-nonconforming-patients. Accessed November 1, 2021.
9. Wilson BDM, Meyer IH: Nonbinary LGBTQ adults in the United States. Available at: https://williamsinstitute.law.ucla.edu/publications/nonbinary-lgbtq-adults-us/. Accessed November 1, 2021.
10. Lippa RA (2000): Gender-related traits in gay men, lesbian women, and heterosexual men and women: The virtual identify of homosexual-heterosexual diagnosticity and gender diagnosticity. J Pers 68:899–926.
11. Marshal MP, Dietz LJ, Friedman MS, Stall R, Smith HA, McGinley J, *et al.* (2011): Suicidality and depression disparities between sexual minority and heterosexual youth: A meta-analytic review. J Adolesc Health 49:115–123.
12. Barr PB, Bigdeli TB, Meyers JL (2022): Prevalence, comorbidity, and sociodemographic correlates of psychiatric disorders reported in the All of Us research program. JAMA Psychiatry 79:622–628.
13. Woodford MR, Weber G, Nicolazzo Z, Hunt R, Kulick A, Coleman T, *et al.* (2018): Depression and attempted suicide among LGBTQ college students: Fostering resilience to the effects of heterosexism and cisgenderism on campus. J Coll Stud Dev 59:421–438.
14. Green AE, DeChants JP, Price MN, Davis CK (2022): Association of gender-affirming hormone therapy with depression, thoughts of suicide, and attempted suicide among transgender and nonbinary youth. J Adolesc Health 70:643–649.
15. Polderman TJC, Kreukels BPC, Irwig MS, Beach L, Chan YM, Derks EM, *et al.* (2018): The biological contributions to gender identity and gender diversity: Bringing data to the table. Behav Genet 48:95–108.
16. Burri A, Cherkas L, Spector T, Rahman Q (2011): Genetic and environmental influences on female sexual orientation, childhood gender typicality and adult gender identity. PLoS One 6:e21982.
17. Theisen JG, Sundaram V, Filchak MS, Chorich LP, Sullivan ME, Knight J, *et al.* (2019): The use of whole exome sequencing in a cohort of transgender individuals to identify rare genetic variants. Sci Rep 9:20099.
18. Sanders AR, Beecham GW, Guo S, Dawood K, Rieger G, Krishnappa RS, *et al.* (2021): Genome-wide linkage and association study of childhood gender nonconformity in males. Arch Sex Behav 50:3377–3383.
19. Abdellaoui A, Verweij KJH (2021): Dissecting polygenic signals from genome-wide association studies on human behaviour. Nat Hum Behav 5:686–694.
20. Ganna A, Verweij KJH, Nivard MG, Maier R, Wedow R, Busch AS, *et al.* (2019): Large-scale GWAS reveals insights into the genetic architecture of same-sex sexual behavior. Science 365:eaat7693.
21. Feliciano P, Daniels AM, Snyder LG, Beaumont A, Camba A, Esler A, *et al.* (2018): SPARK Consortium. SPARK: A US cohort of 50,000 families to accelerate autism research. Neuron 97:488–493.
22. Hisle-Gorman E, Landis CA, Susi A, Schvey NA, Gorman GH, Nylund CM, Klein DA (2019): Gender dysphoria in children with autism spectrum disorder. LGBT Health 6:95–100.
23. Warrier V, Greenberg DM, Weir E, Buckingham C, Smith P, Lai MC, *et al.* (2020): Elevated rates of autism, other neurodevelopmental and psychiatric diagnoses, and autistic traits in transgender and gender-diverse individuals. Nat Commun 11:3959.
24. Achenbach TM, Rescorla L (2003): Manual for the ASEBA Adult Forms and Profiles. Burlington, VT: ASEBA, University of Vermont.
25. Strang JF, Wallace GL, Michaelson JJ, Fischbach AL, Thomas TR, Jack A, *et al.* (2023): The Gender Self-Report: A multidimensional gender characterization tool for gender-diverse and cisgender youth and adults. Am Psychol 78:886–900.
26. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, *et al.* (2015): A global reference for human genetic variation. Nature 526:68–74.
27. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al.* (2007): PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575.
28. Taliun D, Harris DN, Kessler MD, Carlson J, Szpiech ZA, Torres R, *et al.* (2021): Sequencing of 53,831 diverse genomes from the NHLBI TOPMed program. Nature 590:290–299.
29. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, *et al.* (2016): Next-generation genotype imputation service and methods. Nat Genet 48:1284–1287.
30. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, *et al.* (2021): Twelve years of SAMtools and BCFtools. GigaScience 10:giab008.
31. Privé F, Aschard H, Ziyatdinov A, Blum MGB (2018): Efficient analysis of large-scale genome-wide data with two R packages: Bigstatsr and bigsnpr. Bioinformatics 34:2781–2787.
32. Privé F, Luu K, Blum MGB, McGrath JJ, Vilhjálmsson BJ (2020): Efficient toolkit implementing best practices for principal component analysis of population genetic data. Bioinformatics 36:4449–4457.
33. Yang J, Lee SH, Goddard ME, Visscher PM (2011): GCTA: A tool for genome-wide complex trait analysis. Am J Hum Genet 88:76–82.
34. Privé F, Arbel J, Vilhjálmsson BJ (2021): LDpred2: Better, faster, stronger. Bioinformatics 36:5424–5431.
35. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, *et al.* (2019): Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Genet 51:63–75.
36. Watson HJ, Yilmaz Z, Thornton LM, Hübel C, Coleman JRI, Gaspar HA, *et al.* (2019): Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. Nat Genet 51:1207–1214.
37. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, *et al.* (2019): Identification of common genetic risk variants for autism spectrum disorder. Nat Genet 51:431–444.

38. Mullins N, Forstner AJ, O'Connell KS, Coombes B, Coleman JRI, Qiao Z, *et al.* (2021): Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. Nat Genet 53:817–829.
39. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, *et al.* (2019): Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci 22:343–352.
40. International Obsessive Compulsive Disorder Foundation Genetics Collaborative (IOCDF-GC) and OCD Collaborative Genetics Association Studies (OCGAS) (2018): Revealing the complex genetic architecture of obsessive–compulsive disorder using meta-analysis. Mol Psychiatry 23:1181–1188.
41. Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, *et al.* (2022): Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature 604:502–508.
42. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, *et al.* (2018): Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet 50:1112–1121.
43. Becker J, Burik CAP, Goldman G, Wang N, Jayashankar H, Bennett M, *et al.* (2021): Resource profile and user guide of the Polygenic Index Repository. Nat Hum Behav 5:1744–1758.
44. van den Berg SM, de Moor MH, Verweij KJ, Krueger RF, Luciano M, Arias Vasquez A, *et al.* (2016): Meta-analysis of genome-wide association studies for extraversion: Findings from the genetics of personality consortium. Behav Genet 46:170–182.
45. Nagel M, Jansen PR, Stringer S, Watanabe K, de Leeuw CA, Bryois J, *et al.* (2018): Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. Nat Genet 50:920–927.
46. De Moor MH, Costa PT, Terracciano A, Krueger RF, de Geus EJ, Toshiko T, *et al.* (2012): Meta-analysis of genome-wide association studies for personality. Mol Psychiatry 17:337–349.
47. Karlsson Linnér RK, Biroli P, Kong E, Meddens SFW, Wedow R, Fontana MA, *et al.* (2019): Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nat Genet 51:245–257.
48. Neale Laboratory: UK Biobank GWAS – Round 2. Available at: http://www.nealelab.is/uk-biobank/. Accessed October 20, 2021.
49. Champely S: Pwr: Basic functions for power analysis. Available at: https://CRAN.R-project.org/package=pwr. Accessed September 1, 2020.
50. Keller MC (2014): Gene × environment interaction studies have not properly controlled for potential confounders: The problem and the (simple) solution. Biol Psychiatry 75:18–24.
51. Zietsch BP, Sidari MJ, Abdellaoui A, Maier R, Långström N, Guo S, *et al.* (2021): Genomic evidence consistent with antagonistic pleiotropy may help explain the evolutionary maintenance of same-sex sexual behaviour in humans. Nat Hum Behav 5:1251–1258.
52. Clarke TK, Lupton MK, Fernandez-Pujals AM, Starr J, Davies G, Cox S, *et al.* (2016): Common polygenic risk for autism spectrum disorder (ASD) is associated with cognitive ability in the general population. Mol Psychiatry 21:419–425.