Polygenic scores clarify the relationship between mental health and gender diversity

Taylor R. Thomas[1], Ashton J. Tener[1], Amy M. Pearlman [2], Katherine L. Imborek [3], Ji Seung Yang[4], John F. Strang[5], Jacob J. Michaelson[1,6,7*]

1 Department of Psychiatry, University of Iowa, Iowa City, IA
2 Department of Urology, University of Iowa, Iowa City, IA
3 Department of Family Medicine, University of Iowa, Iowa City, IA
4 Department of Human Development and Quantitative Methodology, University of Maryland, College Park, MD
5 Gender and Autism Program, Center for Neuroscience, Children's National Hospital, George Washington University School of Medicine, Washington DC
6 Iowa Neuroscience Institute, University of Iowa, Iowa City, IA
7 Hawkeye Intellectual and Developmental Disabilities Research Center (Hawk-IDDRC), University of Iowa, Iowa City, IA
***** jacob-michaelson@uiowa.edu

Abstract

Gender diverse individuals are at higher risk for mental health problems. What remains unclear is whether this increased risk is attributable to environmental stressors (e.g., minority stress), innate pleiotropic effects on gender diversity and mental health, or gene-by-environment interactions. Here, we present two studies (N = 5,388 and N = 696, all from the SPARK autism cohort) investigating these questions. First, N = 696 independent adults (58% autistic) provided thorough information on their gender identities and expression using the Gender Self Report (GSR), a novel assessment for the continuous, multidimensional characterization of gender diversity. We tested 20 behavioral polygenic scores for association with the two GSR dimensions: Binary Gender Diversity (degree of identification with the gender opposite that implied by sex designated at birth) and Nonbinary Gender Diversity (degree of identification with a gender that is neither man/male nor woman/female). We found no evidence of association between gender diversity and polygenic risk for adult-onset psychiatric conditions (major depression, bipolar disorder, schizophrenia). Strikingly, we instead found that both gender diversity dimensions were positively associated with the cognitive performance polygenic score (Binary $\rho = 0.09,\ p = 0.02$ and Nonbinary $\rho = 0.11,\ p = 0.003$). We also found Binary Gender Diversity to be positively associated with polygenic scores for autism ($\rho = 0.08,\ p = 0.045$) and non-heterosexual sexual behavior ($\rho = 0.09,\ p = 0.02$). Additionally, we tested for interactions effects and found no association between increasing gender diversity and poorer mental health outcomes in the subsample with low genetic risk for these neuropsychiatric conditions. Only in the subsample with high genetic risk for major depression or schizophrenia did we observe a significant relationship between gender diversity and poor mental health outcomes. In the second study (N = 5,388), we used self-reported categorical gender identity and found that, in agreement with our first study, transgender and gender nonbinary individuals had significantly higher polygenic scores for cognitive performance ($t = 3.90,\ p = 0.0001$) and non-heterosexual sexual behavior ($t = 2.0,\ p = 0.04$), as well as risky behavior ($t = 2.10,\ p = 0.04$). In summary, our findings demonstrate gender diversity to have pleiotropic relationships with adaptive traits such as cognitive performance, implicate the environment (e.g., minority stress) as a key factor interacting with polygenic risk to generate poor mental health outcomes, and challenge a pathologizing view of gender diversity.

1 Introduction

Sex and gender (see Table 1 for our definitions of terms) have major impacts on health [1]. This stems from both extrinsic factors (e.g., healthcare barriers [2, 3]) as well as biological factors, with sex and gender modulating the underlying molecular mechanisms of disease and well-being [4]. In health research, sex has been a more objective and well-defined variable than gender, which is multidimensional with binary and nonbinary components and often experienced on a continuum [5]. Gender diversity can be reported through self-endorsement of gender identity labels (e.g., transgender, nonbinary, genderqueer, demi-boy), but these labels are contextually and culturally dependent (i.e., not accessible by all) and variable and often non-specific in their meanings [6]. Further, there are numerous gender identity self-descriptors, and group-based analyses based on parsing datasets into individual descriptors erode statistical power for meaningful comparisons given the numerous individual subgroupings. Moreover, gender diversity, a fundamental aspect of human diversity, is not only expressed by individuals with transgender and/or gender

nonbinary identities (TGNB). People who identify as cisgender also exhibit some variation in gender diversity that would be lost in studies only reporting categorical descriptors of gender identity [7]. Therefore, a multidimensional, continuous characterization of gender that uses simple and broadly accessible language will enable health researchers to appropriately incorporate gender diversity in their analyses.

Gender diversity is a crucial variable to include in health research, and this may be particularly true in mental health and neuropsychiatric research. Groups that express higher levels of gender diversity than the cisgender proportional majority, such as LGBTQ+ (lesbian, gay, bisexual, transgender, and queer) individuals [8, 9], often have greater rates of anxiety and depression and are more likely to attempt suicide [10]. A recent report of N = 329,038 participants in the All of Us cohort found that the non-heterosexual participants had greater prevalence of all neuropsychiatric diagnoses compared to the heterosexual participants [11]. The exact mechanisms for this are not entirely known. Research has shown that poorer mental health is due to factors related to the experienced adversity from sexual orientation and/or gender diversity minority stress; for example, discrimination and resilience partially mediate negative mental health outcomes in LGBTQ+ college students [12]. Additionally, access to gender-affirming hormone therapy for TGNB youth is associated with a reduced likelihood of depression and suicidality [13]. However, to the best of our knowledge, no study has leveraged genetic data to elucidate the relationships between gender diversity and mental health, so any possible contributions of genetic factors are unknown.

The brain is the biological seat of personal identity, including gender identity. We hypothesize that gender identity is therefore susceptible to genetic influences like other human behavior traits [14]. Most behaviors are somewhat heritable, with genome-wide association studies (GWASs) of common genetic variants showing many loci, each of a small effect, contributing additively (i.e., polygenicity) [15]. Additionally, genomic loci associated with one behavior trait are often found to be associated with another trait, suggesting the two traits have a degree of pleiotropy. One method to estimate pleiotropy is to use polygenic scores that are the genome-wide cumulative sum from a GWAS; a polygenic score is then correlated with the other trait of interest. Genetic research of gender diversity has been limited and underpowered for gene discovery [16, 17].

Among the current well-powered GWAS, the most reasonable proxy to gender diversity is the non-heterosexual sexual behavior GWAS [18] performed in N = 408,995 UK Biobank participants. The trait was defined as the yes/no response to ever having sex with someone of the same sex (the nuance between same-sex versus same-gender are lost due to the nature of the question). The heritability of non-heterosexual sexual behavior varied by age, ranging from 0.08 to 0.25 and was positively genetically correlated with several neuropsychiatric conditions and personality traits. However, the interpretation of these genetic correlations is limited because of the confounding with experienced adversity, meaning the positive correlation could be due to either individuals engaging in non-heterosexual sexual behavior facing more sexual and/or gender-based discrimination that increase risk for neuropsychiatric conditions and/or pleiotropy between non-heterosexual sexual behavior and neuropsychiatric risk. Recent work has begun disentangling the confounding variables of discrimination, genetic risk, and mental health outcomes in a study of N = 1,146 participants. They regressed out the effects of anxiety, depression, and neuroticism polygenic scores from both their discrimination measures (not necessarily sexual or gender-based discrimination) and anxiety measures and found the association between discrimination and anxiety was persistent after controlling for these genetic liabilities [19].

In this study, we investigated whether gender diversity, like non-heterosexual sexual behavior, is pleiotropic with other behavioral traits and how this pleiotropy might play a role in mental health. Study participants were from the SPARK cohort [20], a nationwide genetic study of over 300,000 participants with and without autism. Existing research demonstrating the common intersection of autism and gender diversity makes SPARK an ideal cohort for gender diversity studies. Previous studies have shown there is an enrichment of gender diversity in autistic samples compared to the general population [21]. Likewise, general population samples of transgender and more broadly gender diverse people are more likely to be or autistic or have clinically relevant levels of autistic traits [22]. In our sample of N = 696 participants, we calculated polygenic scores for 20 traits including cognitive ability, personality, and neuropsychiatric conditions and administered two psychometric self-report tools. The first, the Adult Self Report (ASR) [23], is a well-established instrument that measures several mental health outcomes and adaptive behaviors. The second, the Gender Self-Report captures two quantitative dimensions of gender diversity: Binary Gender Diversity, the extent one experiences themselves as the other binary gender (i.e., different from their sex designated at birth), and Nonbinary Gender Diversity, the extent one experiences themselves as not female or male. We then sought to answer the following questions: First, are behavior polygenic scores correlated with the two measures of dimensional gender diversity from the GSR? Second, do the polygenic findings broadly replicate in a larger sample with a categorical gender identity phenotype? Third, is self-reported mental health correlated with gender diversity? Lastly, do polygenic scores provide additional context in our understanding of the relationship between gender diversity and mental health? An overview of our analyses is shown in Figure 1.

2 Results

2.1 Gender diversity and mental health correlations

The demographic characteristics of the SPARK Research Match participants are shown in Table 2. The final sample size was N = 696 participants, with approximately one-third of the cohort identifying as transgender or gender nonbinary (TGNB). Fifty-eight percent of participants were autistic, and 22% were male. The genetic ancestry categorization is based on the five continental populations described by 1000 Genomes [24]. Ninety-two percent of participants were in the Europe genetic group, 4% in the Americas group, 4% in the South Asia group, 1% in the East Asia group, 0% in the Africa group.

The two gender diversity values, Binary and Nonbinary Gender Diversity, are from the Gender Self Report (GSR) [25]. The GSR values range from 0 (no gender diversity) to 1 (high gender diversity), with the mode being near 0 for both; these values were then controlled for age, sex designated at birth, and autism diagnostic status by linear regression, and then standardized to a mean of 0 and a standard deviation of 1. The distributions of these two values (controlling for covariates) are shown in Figure 2A and are colored by self-endorsed sexual orientation (top panel) and gender identity (bottom panel). The overall trend shows higher gender diversity in LGBQ+ and TGNB participants. The GSR values were significantly positively correlated with each other: $\rho = 0.57,\ p < 5 \times 10^{-5}$ (Figure 2B).

The two mental health outcome values, Externalizing and Internalizing, are from the Adult Self Report (ASR) [23]. Internalizing problems is a composite score of anxiety, depression, and somatic complaints, and Externalizing problems is a composite score of aggressive, rule-breaking, and intrusive behavior. These values were also controlled for age, sex designated at birth, and autism diagnostic status by linear regression, and then standardized to a mean of 0 and a standard deviation of 1. Externalizing and Internalizing were significantly positively correlated with each other: $\rho = 0.61,\ p < 5 \times 10^{-5}$ (Figure 2C). The ASR values were also significantly positively correlated with the GSR values (Figure 2D). Binary Gender Diversity was more correlated with Internalizing ($\rho = 0.14,\ p = 0.0001$) than Externalizing ($\rho = 0.10,\ p = 0.01$). Nonbinary Gender Diversity was also more correlated with Internalizing ($\rho = 0.18,\ p < 5 \times 10^{-5}$) than Externalizing ($\rho = 0.13,\ p = 0.0009$).

2.2 Polygenic score correlations with gender diversity and mental health

We next assessed the relationships between the GSR and ASR with twenty polygenic scores for behavior traits spanning across four behavior domains [15]. The first domain is reflective of traits related to cognition and socioeconomic status— cognitive performance and educational attainment [26]. The second domain is personality and well-being traits, with three traits from the Big Five personality (extraversion [27], neuroticism [28], openness [29]), as well as depressive symptoms [30], loneliness [31], risky behavior [32], and subjective well-being (SWB) [33]. The third domain is sexuality and reproduction-related traits: these include age at first birth (i.e., age at parenthood) [34], number of children ever born (NEB) [31], and non-heterosexual sexual behavior [18]. The last domain is neuropsychiatric conditions— ADHD [35], anorexia [36], autism [37], bipolar disorder [38], major depression [39], OCD [40], and schizophrenia [41]. Polygenic scores were controlled for the effects of the first twenty genetic principal components (to account for genetic ancestry effects), as well as age, sex designated at birth, and autism diagnostic status by linear regression. We computed correlation coefficients between the polygenic scores with the GSR values (Figure 3A) and ASR values (Figure 3B).

As expected, the non-heterosexual sexual behavior polygenic score was significantly positively correlated with Binary Gender Diversity: $\rho = 0.09,\ p = 0.02$. The non-heterosexual sexual behavior polygenic score was also positively correlated with Nonbinary Gender Diversity, although the correlation did not reach nominal significance: $\rho = 0.05,\ p = 0.16$. Strikingly, the cognitive performance polygenic score was significantly positively correlated with Binary Gender Diversity ($\rho = 0.09,\ p = 0.01$) and Nonbinary Gender Diversity ($\rho = 0.11,\ p = 0.003$), meaning that polygenic propensity for greater cognitive performance was associated with elevated binary and nonbinary gender diversity. The autism polygenic score was significantly positively correlated with Binary Gender Diversity: $\rho = 0.08,\ p = 0.04$. No other neuropsychiatric polygenic scores were significantly correlated with the GSR values. However, we did observe significant neuropsychiatric polygenic score correlations with the two ASR values. Externalizing was positively correlated with the ADHD polygenic score ($\rho = 0.13,\ p = 0.0008$) and negatively correlated with the anorexia polygenic score ($\rho = -0.08,\ p < 0.05$) and the openness polygenic score ($\rho = -0.08,\ p = 0.04$). Internalizing was positively correlated with the depression polygenic score ($\rho = 0.10,\ p = 0.01$), as well as the depressive symptoms polygenic score ($\rho = 0.10,\ p = 0.007$) and the neuroticism polygenic score ($\rho = 0.12,\ p = 0.002$).

We performed the same correlations stratified by autism diagnostic status, with the results being comparable (Figure S1). Results were also comparable when performing the correlations only in the European genetic population group of N = 639 (Figure S2).

2.3 Replication of polygenic score differences for categorical gender identity

We next tested if the polygenic score associations we identified using dimensional gender diversity measures were similar in a larger sample (N = 5,388) with a categorical gender identity phenotype. For this study, we used the basic background history files to label individuals as either cisgender or TGNB. Specifically, we used discordance between the participant's designated sex at birth (options: Male or Female) and their gender (options: Male, Female, or Other) to categorize if the participant was TGNB or cisgender. We also merged this dataset with the Research Match for a final sample size of N = 5,388, with N = 590 from the Research Match and the remaining N = 4,798 from the background history reporting. The demographic summary of this larger cohort is also shown in Table 2. The average age was 25 years old, and 88% of participants were autistic.

We tested for polygenic score differences between the two gender identity groups: cisgender (N = 5,089) versus TGNB (N = 299) with the results shown in Figure 4. The strongest effect was seen for the cognitive performance polygenic score, with the TGNB group being significantly higher ($N = 5,388,\ x_d = 0.24[0.12 - 0.36]$, $t = 3.90,\ p = 0.0001$). The TGNB group also had significantly higher polygenic scores for risky behavior ($x_d = 0.12[0.01 - 0.23]$, $t = 2.10,\ p = 0.04$) and non-heterosexual sexual behavior ($x_d = 0.12[0 - 0.23]$, $t = 2.0,\ p = 0.04$).

2.4 Interactions between gender diversity, mental health, and polygenic scores

In order to investigate whether polygenic risk and gender diversity interact in modeling mental health outcomes, we tested for interaction effects in linear models and also performed stratified correlations. We grouped participants into one of three groups for each polygenic score: high risk (upper quartile, N = 174), neutral (2nd and 3rd quartiles, N = 348), and the low risk (lower quartile, N = 174). We compared GSR-ASR associations between the polygenic score upper group (coded as 1) versus the polygenic score lower group (coded as 0) and removed the neutral risk group. We first formally tested for polygenic group-by-GSR interaction effects with the linear model `ASR value ~ GSR value + polygenic group + GSR value:polygenic group`. The interaction terms are shown in Figure 5A, with nominally significant interactions ($p < 0.05$) indicated by a white asterisk. We then performed GSR-ASR correlations stratified by the polygenic group, and Figure 5B shows the ρ for GSR-ASR correlations for the upper quartile versus lower quartile polygenic risk groups. Figure 5C shows the stratified correlations for the two strongest polygenic group-by-GSR interaction effects.

We identified four significant polygenic group-by-GSR interactions, specifically the schizophrenia and depression polygenic risk. Within the entire cohort of N = 696, Nonbinary Gender Diversity and Internalizing are positively correlated: $\rho = 0.18,\ p < 5 \times 10^{-5}$. However, this apparent main effect appears to be driven by a context-specific interaction with genetic risk: in the subset at greatest schizophrenia polygenic risk (e.g. the upper quartile, N = 174), the correlation between Nonbinary Gender Diversity and Internalizing is $\rho = 0.36,\ p < 5 \times 10^{-5}$. While in the lower risk group (e.g. the lower quartile, N = 174), there is no correlation: $\rho = 0.05,\ p = 0.52$. The effect is similar when stratifying by the depression polygenic score— the high risk correlation is $\rho = 0.30,\ p = 0.00006$, while in the low risk group the correlation is not significant: $\rho = 0.09,\ p = 0.22$.

3 Discussion

Our analyses are the first to address the relationships of multidimensional gender diversity with mental health and genetics. We leveraged two novel, quantitative measures of gender diversity, Binary and Nonbinary Gender Diversity, from the Gender Self-Report (GSR) in a neurodiverse sample of N = 701 adults participating in the SPARK autism study. In this sample, we found greater gender diversity in female, autistic, and LGBTQ+ participants. Due to the structure of SPARK and study recruitment, we were only able to collect data from independent adults with autism or immediate family-members of someone with autism (mostly parents). Therefore, the elevated gender diversity in the autistic subset should be interpreted with the caveat that the non-autistic participants were older and presumed to adhere to more traditional gender roles. Still, these results are in line with prior research that has shown the enrichment for gender diversity in autism [21]. Intriguingly, while our results showed higher gender diversity in the LGBTQ+ participants, many people who identify as cisgender also showed evidence of gender diversity, though not enough for them to report being transgender or more broadly gender diverse. This underscores the value of the GSR in capturing dimensional gender diversity beyond self-endorsed identities, alone. The formation of gender identity is a complex and multi-factorial process [42] and is contextualized by numerous factors like time (e.g., age, generation), region, and culture. Additionally, the conceptualization of these identities requires understanding of how the self relates to other points of reference. This can be different for some autistic people who may struggle with understanding social and gender norms [43].

We correlated 20 behavior polygenic scores with the two GSR measures, and strikingly, the strongest association was cognitive performance being positively associated with both Binary and Nonbinary Gender Diversity (Figure 3A). This finding was validated in our larger sample of N = 5,388 with categorical gender identity (cisgender versus TGNB); the cognitive performance polygenic score was significantly higher in the TGNB group compared to the cisgender group (Figure 4). This suggests cognitive capacity may be an important component in the development of more complex and nuanced gender identities. Beyond cognitive performance, we also found the non-heterosexual sexual behavior polygenic score to be positively correlated with Binary Gender Diversity. While gender identity and sexual orientation are distinct concepts, the non-heterosexual sexual behavior genome-wide association study (GWAS) is the most well-powered GWAS that is adjacent to gender diversity. Non-heterosexual behavior is associated with reduced number of children (i.e., reduced reproductive fitness) [44], so the population endurance of alleles associated with non-heterosexual behavior is an interesting conflict. Among heterosexuals, the non-heterosexual sexual behavior polygenic score was recently shown to be positively correlated with an increased number of partners, which presumably increases reproductive fitness [45]. Building off this, our results suggest gender diversity may part of a pleiotropic ensemble of traits with adaptive advantages (e.g., cognitive performance).

We expected neuropsychiatric polygenic scores to also be positively correlated with the GSR measures, considering non-heterosexual sexual behavior shows positive genetic correlation with several neuropsychiatric conditions [18]. In light of this prior research, it was surprising that we found no significant positive correlations with GSR values and neuropsychiatric polygenic scores, aside from Binary Gender Diversity being positively correlated with the autism polygenic score. This suggests that, within the statistical power limits of our sample, gender diversity is not in strong pleiotropic relationships with adult-onset psychiatric disorders. Instead, in our sample greater gender diversity appears to have pleiotropic relationships with higher cognitive ability, non-heterosexual sexual behavior, and autism.

The lack of a genetic main effect linking psychiatric conditions and gender diversity, combined with our observation that the GSR values nevertheless show numerous significant correlations with poorer self-reported mental health (Figure 2D) prompted us to examine the possibility of a relationship between gender diversity and mental health that depends on genetic risk level (i.e., an interaction between polygenic risk and gender diversity). To accomplish this, we used the polygenic score for each psychiatric condition to stratify our sample into high and low risk groups (upper and lower quartiles of polygenic scores, respectively, each with N=174, see Figure 5A). We observed dramatic differences in the correlations between the genetic risk groups when stratifying by the schizophrenia and depression polygenic scores (Figure 5B, C): the groups of high depression and schizophrenia polygenic risk had the strongest GSR-ASR correlations, whereas the correlations in the low-risk groups were absent (i.e., not nominally significant). This suggests that polygenic risk for depression and schizophrenia interact with gender diversity (or environmental factors related to gender diversity such as discrimination and/or minority stress) in determining mental health outcomes. In other words, our findings provide evidence that the robustly observed relationship between gender diversity and mental health outcomes is not solely environmental or genetic, but rather a combination of the two. Specifically, an individual's polygenic risk for psychiatric disorders determines the extent their gender diversity (and/or experiences of adversity that gender diverse individuals may experience) impacts their mental health. This observation could also be cast in terms of resilience: the high genetic risk group is less resilient against experienced adversity that might impact mental health, while the low risk group shows more resilience against poorer mental health as gender diversity and/or associated stressors increase. This interpretation is congruent with previous work that found that the individuals at high polygenic risk for depression were more likely to have more depressive symptoms while under stress, and those in the lowest depression polygenic risk group were least likely/most resilient under stress [46].

Our results and their interpretations have several limitations. Most genetic analyses (genome-wide association studies, heritabilities, polygenic scores) require large sample sizes due to the small effects of individual common variants. Consequently, our primary limitation is the small sample size, and we therefore were only powered to detect strong polygenic score correlations. With our sample size of N = 696, we were at 80% power to detect correlations greater than $\rho = \pm 0.106$. Additionally, age, sex designated at birth, and autism diagnostic status are entangled with other variables of interest. Autism diagnosis is confounded at the genetic level, as observed in previous work that showed that educational attainment [37] and cognitive performance [47] are positively genetically correlated with autism. However, we repeated our analyses stratified by autism diagnostic status and found the results to be comparable (Figure S1). Future work with larger samples should analyze the interplay between autism, sex designated at birth, and polygenic scores in their associations with gender diversity by performing sufficiently powered analyses stratified by autism and designated sex.

In summary, our findings show that gender diversity, as captured by the Gender Self Report, has dimensional properties that share common genetic factors with cognitive performance, non-heterosexual sexual behavior, and autism. In agreement with previous studies, we find greater gender diversity to be correlated with poorer mental health, but this relationship is not due to shared genetic effects between psychiatric disorders and gender diversity. Rather, one's polygenic background is a risk/resilience mechanism that interacts with gender diversity (and/or the

adversity that comes with it) in determining mental health outcomes.

4 Materials and methods

4.1 Sample description

SPARK [20] is a U.S.-based nationwide autism study of over 300,000 participants, with genetic data available for many of the participants. SPARK is approved by the Western IRB (#20151664). Independent adults, with or without autism, were invited to participate in our SPARK Research Match. The Research Match study was approved by the University of Iowa Institutional Review Board (IRB #201611784). Those who agreed and consented to participate were asked to complete the Gender Self Report (GSR) [25], the Adult Self Report (ASR) [23], and additional questions regarding their sexual orientation, gender identity, and gender expression, with the final sample size N = 818. We then removed nine individuals who had withdrawn from the SPARK study since the Research Match based on the Version 8 phenotype release. N = 696 is the final sample size after genetic data availability and quality control filtering.

For the larger cohort, we used the Version 8 independent adult, child, and sibling background history files. The independent adult dataset was self-report, whereas the child and sibling datasets were parent-report. We only kept children age 14 or above and whose cognitive impairment status at enrollment was not significantly below age.

4.2 Research Match measures

Self-endorsed labels of gender identity and sexual orientation: Participants were able to select as many labels for gender identity and sexual orientation 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 selecting lesbian, gay, bisexual, pansexual, homosexual, queer, and/or polysexual were grouped as LGBQ+ and heterosexual orientation was categorized 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 (GSR): The Gender Self-Report (GSR) itemset was developed through an iterative multi-input community driven process with autistic cisgender, autistic gender-diverse, and non-autistic cisgender and gender-diverse collaborators [25]; Open Science Framework Development Summary: `https://osf.io/qh25d/?view_only=c0ce41d07bca4af1b792e074d51b7ded`. A diversified recruitment approach was employed across seven separate recruitments (N = 1,654), including the current study's recruitment (N = 818), to optimize the breadth of the GSR calibration sample and enrich the sample based on the following key characteristics: autism, gender-diverse identities (binary and nonbinary), the intersection of autism and gender-diverse identities, transition age/young adult age, and female designation at birth within the entire sample and within autism, specifically. This sampling approach resulted in an overall calibration sample that was 37.5% autistic, 32.6% gender diverse, and 38.9% cisgender sexual minority. Two-dimensional graded response model with a normal-mixture latent density adequately fit the data and yielded two factors. The two factors are labeled Female-Male Continuum and Nonbinary Gender Diversity. A transformation of the Female-Male Continuum values based on designated sex at birth produced Binary Gender Diversity values (i.e., representing the distance on the binary gender spectrum from individual's designated sex at birth). GSR calibration employed differential item functioning, an equity-based psychometric method to identify and reduce bias, in this case by age as well as autism status. Empirical reliability coefficients for response pattern expected a posteriori scores were 0.75 for Nonbinary Gender Diversity and 0.85 for Binary Gender Diversity. GSR factors performed well across the following validation metrics: (1) construct validity; GSR factor values followed expected value patterns comparing gender identity subgroups, (2) convergent validity; GSR factor values correlated with existing gender-related measures and in expected directions, and (3) ecological validity; GSR factor values aligned with report of gender-affirming medical treatment request/receipt. The final GSR itemset is composed of 30 questions, that participants answered: 1 = never true, 2 = sometimes true, 3 = often true, 4 = always true. In our genetic sample of N = 696 participants, these two GSR values were controlled for age in months, sex designated at birth, and autism diagnostic status by linear regression and then standardized to a mean of 0 and a standard deviation of 1. These values were then used as the phenotypes in the subsequent correlation analyses.

Adult Self Report (ASR): The Adult Self Report (ASR) [23] is a well-established self-report questionnaire of 129 items assessing a range of adaptive behaviors and mental health outcomes. The participants responds either: 0 = not true, 1 = somewhat or sometimes true, or 2 = very true or often true. From the N = 809, five participants were removed due to having 12 (approximately 10%) or more missing ASR items. In the remaining N = 809, 0.2% of the data was missing, with no item having more than five missing data points. The missing data was imputed to the

median. The two measures used in our analyses were Internalizing and Externalizing problems that are summed syndrome subscales. Externalizing problems is composed of aggression, rule-breaking, and intrusive behavior subscales (35 items total), and Internalizing problems is composed of the anxiety, depression, and somatic complaints subscales (37 items total). In our genetic sample of N = 696 participants, these two ASR values were controlled for age, sex designated at birth, and autism diagnostic status by linear regression and then standardized to a mean of 0 and a standard deviation of 1. These values were then used as the phenotypes in the subsequent correlation analyses.

4.3 Categorical gender identity measures for larger cohort

If the participant's designated sex at birth (options: Male or Female) did not match their gender (options: Male, Female, or Other) then the participant was categorized as TGNB. We then merged with our Research Match gender identity labels. This merged cohort was then filtered to individuals in which genetic data was available that passed quality control (see next section for genotype quality control) and were not genetically related by a relatedness threshold of 0.125. The final sample size was N = 5,388 with N = 590 from the Research Match and the other N = 4,798 from the background history files.

4.4 Genotype quality control and imputation

We used the genotype array data from SPARK integrated whole-exome-sequencing (iWES1) 2022 Release and the SPARK whole-genome-sequencing (WGS) Release 2, 3, and 4. iWES1 (N = 69,592) was quality controlled on release, including removing samples due to heterozygosity or high missingness, so no further quality control was performed by us before genotype imputation. iWES1 also provided genetic ancestry assignments based on the 1000 Genomes populations [24]. WGS Release 2 (N = 2,365), Release 3 (N = 2,871), and Release 4 (N = 3,684) were not quality controlled on release, so we performed quality control using PLINK [48] before genotype imputation. First, we removed participants from the WGS releases if they were in iWES1. Second, we removed variants with missingness greater than 0.1 and participants with missingness greater than 0.2. Third, we merged the three releases and then removed any participant whose heterozygosity (F statistic) was not within 3 standard deviations of the mean heterozygosity across the three releases. We then used the TopMed reference panel [49] to identify strand flips. The final sample size for WGS 2-4 was N = 8,152. iWES1 and WGS 2-4 were then imputed to the TopMed [49] reference panel using the Michigan Imputation Server [50] with the phasing and quality control steps included and to output variants with imputation quality r2 > 0.3. After imputation, the variants were filtered to only the HapMap SNPs (N = 1,054,330 variants) with imputation quality r2 > 0.8 using bcftools [51]. They were lifted over from hg38 to hg19 using the VCF-liftover tool (`https://github.com/hmgu-itg/VCF-liftover`) and the alleles normalized to the hg19 reference genome. Finally, the files were merged and only variants with 0% missingness were retained (N = 914,328).

4.5 Genetic ancestry

Genetic principal components (PCs) were calculated using the bigsnpr package [52], specifically by following the author's recommendations [53] and their 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 the KING threshold of $2^{-3.5}$, 2.) performed principal component analysis using the `bed_autoSVD` on just the unrelated participants, 3.) detected principal component outliers and removed them, 4.) recalculated the principal components, and 5.) projected the principal components onto the entire cohort using the `bed_projectSelfPCA` function. We chose to not remove participants based on their genetic ancestry and instead use genetic ancestry as a continuous variable (instead of categorical), as per recent recommendations [54]. However to establish faith in the robustness of our results, we used the top 40 principal components and performed k-means clustering with K = 5 (for the five populations from 1000 Genomes [24]) and used the genetic ancestry labels from iWES1 to assign labels to the genetic population clusters. We then repeated the polygenic score correlations in the European subset, with the results provided in the Supplemental Information.

4.6 Relatedness

From the N = 809 Research Match participants whom completed the GSR, N = 804 also had sufficient ASR data, and N = 727 had genetic data. This subset of N = 727 was pruned to remove related participants using GCTA [55] with a relatedness threshold of 0.125, corresponding to approximately third degree relatives (N = 31 individual removed).

For the analyses with the larger cohort, we first retained only one participant from each family based on family ID, with prioritization towards TGNB identity. We then removed related participants using GCTA [55] with the relatedness threshold of 0.125. After this filter, N = 5,388 participants remained.

4.7 Polygenic score calculations

Polygenic scores were calculated using LDpred2 [56] and the bigsnpr tools [52] in R [57]. Because SPARK is family-based, an external LD reference based on 362,320 individuals in UK Biobank (provided by the authors of LDpred2) was used to calculate the genetic correlation matrix, estimate heritability, and calculate the infinitesimal beta weights. Polygenic scores were calculated from the following genome-wide association studies performed by the Psychiatric Genomics Consortium: ADHD (2019) [35], anorexia nervosa (2019) [36], autism (2019) [37], bipolar disorder (2021) [38], major depression (2019) [39], OCD (2018) [40], and schizophrenia (2020) [41]. Polygenic scores were calculated from genome-wide association studies performed by the Social Science Genetic Association Consortium for cognitive performance (2018) and educational attainment (2018) [26] and from the UK Biobank for non-heterosexual sexual behavior [18]. The public LDpred2 beta weights from the Polygenic Index Repository [58] were used to calculate polygenic scores for depressive symptoms [30], extraversion [27], loneliness [31], neuroticism [28], openness [29], risky behavior [32], subjective well-being [33], age at first birth [34], number of children ever born (men) [31], and number of children ever born (women) [31].

To control for genetic ancestry confounding with the polygenic scores, we residualized using linear regression for the first 20 genetic principal components. We also additionally controlled for the effects of age in months, sex designated at birth, and autism diagnostic status by linear regression. Lastly, the polygenic scores were standardized to a mean of 0 and a standard deviation of 1. We performed this polygenic scores processing separately for the Research Match cohort (N = 696) and the larger cohort (N = 5,388).

4.8 Polygenic score analyses

Polygenic scores were correlated with the two GSR values (Binary and Nonbinary) and the two ASR values (Externalizing and Internalizing) using Spearman correlations. In the correlations stratified by polygenic risk, we grouped people into one of three groups for each polygenic score: upper 75th quartile (N = 175), middle quartile (N = 351), and the lower 25th quartile (N = 175). We then compared Spearman correlation coefficients between the upper versus lower quartiles. We tested for polygenic group-by-GSR interaction effects in association with the ASR values with linear models: `ASR value ~ GSR value + polygenic group + GSR value:polygenic group`. We used the `pwr.r.test()` function from the R pwr package [59] to determine statistical power for the correlations.

For the larger cohort analyses, two-sample t-tests were used.

Acknowledgments

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

Data and code availability

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

Funding

This work was supported by the National Institutes of Health (MH105527 and DC014489 to JJM) and the National Institute of Mental Health (R01MH100028 to JFS), as well as grants from the Simons Foundation (SFARI 516716 to JJM), the Clinical and Translational Science Award (KL2TR001877 to JJM), the Fahs-Beck Fellow Grant to JFS, and the National Institutes of Health Predoctoral training grant (T32GM008629 to TRT) This work was supported by the University of Iowa Hawkeye Intellectual and Developmental Disabilities Research Center (Hawk-IDDRC) through the Eunice Kennedy Shriver National Institute of Child Health and Human Development (P50HD103556).

Conflicts of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author contributions

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

Figures and Tables

Table 1. Working definitions of terms.
Some definitions are from [7] and [60].

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.
transgender and gender nonbinary (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
neurodivergent	individuals with cognitive differences from the neuromajority in the areas of social, thinking, and/or other neurocognitive process; autism and ADHD are often considered forms of neurodivergence

Table 2. Demographics. The first study used the two continuous measures of gender diversity from the Gender Self Report, and the second study used categorical gender identity.

	1st study		2nd study	
variable	N	%	N	%
total sample size	696	-	5,388	-
age (mean)	37	-	25	-
male	153	22%	3,271	61%
autistic	403	58%	4,731	88%
cisgender identity	424	61%	5,089	94%
TGNB identity	206	30%	299	6%
no gender identity labels	66	9%	-	-
overlap with Research Match	-	-	590	11%
Africa genetic pop'l	0	0%	0	0%
Americas genetic pop'l	27	4%	100	2%
East Asia genetic pop'l	4	1%	4	0%
Europe genetic pop'l	639	92%	5,165	96%
South Asia genetic pop'l	26	4%	119	2%

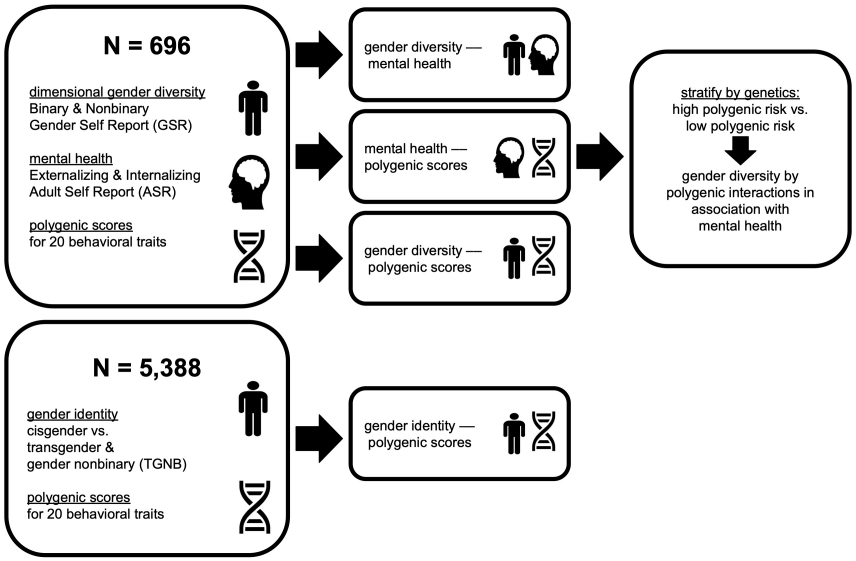

Figure 1. Overview of the study.

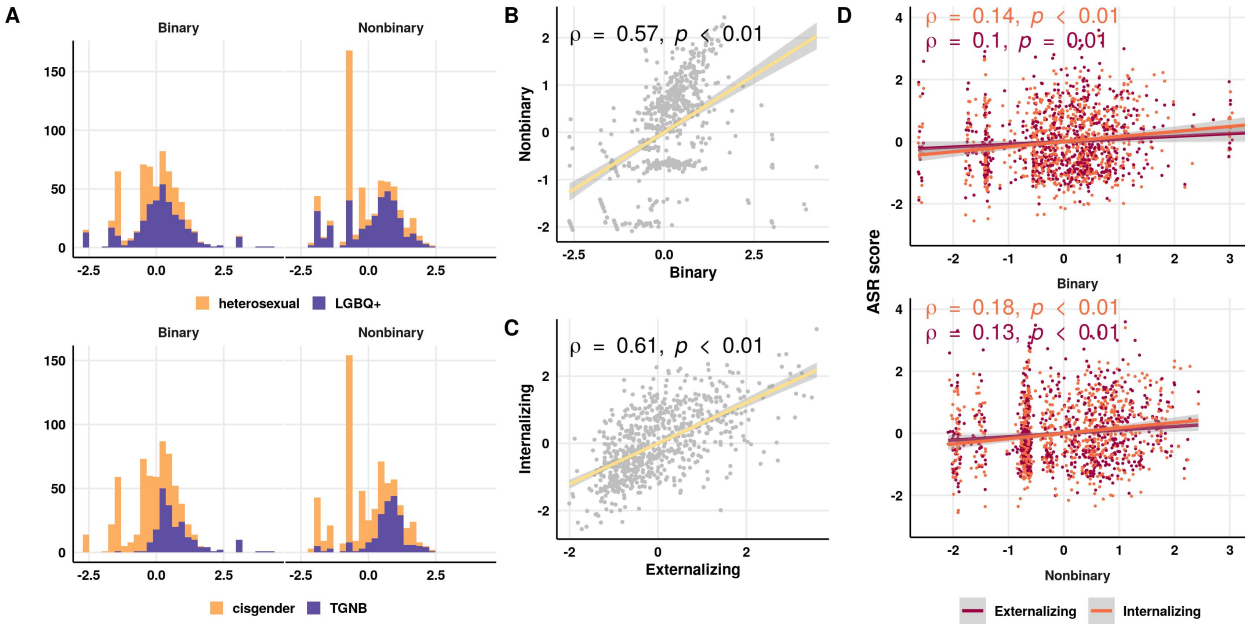

Figure 2. Distributions and correlations of the gender diversity and mental health measures
A Distribution of the two dimensional gender diversity measures from the Gender Self Report (GSR): Binary and Nonbinary Gender Diversity. The GSR values are corrected for age, sex designated at birth, and autism diagnosis, and the histogram are colored by self-endorsed sexual orientation labels (top panel) and gender identity labels (bottom panel). **B** Correlation of the two GSR values. **C** Correlation of the two mental heath measures from the Adult Self Report (ASR): Internalizing and Externalizing. **D**: Correlations of the GSR values with the ASR values.

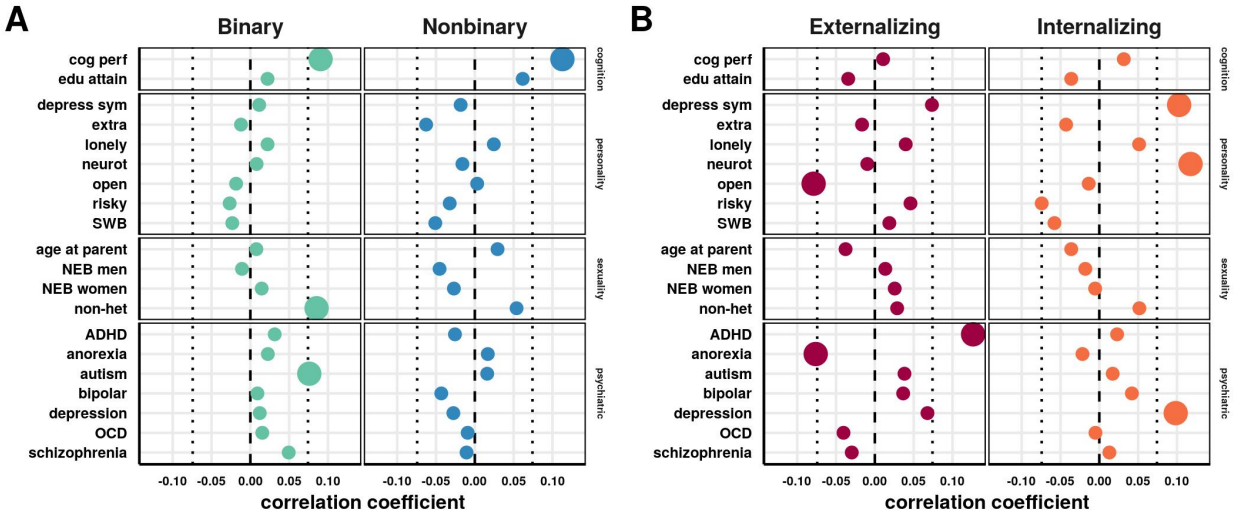

Figure 3. Polygenic score correlations with gender diversity and mental health
Polygenic score correlations with the GSR values **A** and ASR values **B** . The dotted lines represent the minimal correlation coefficient for statistical significance (nominal $p < 0.05$) that is $\rho = \pm 0.074$. Polygenic score abbreviations: extra = extraversion, neurot = neuroticism, open = openness, risky = risky behavior, SWB = subjective well-being, NEB = number of children ever born, non-het = non-heterosexual sexual behavior.

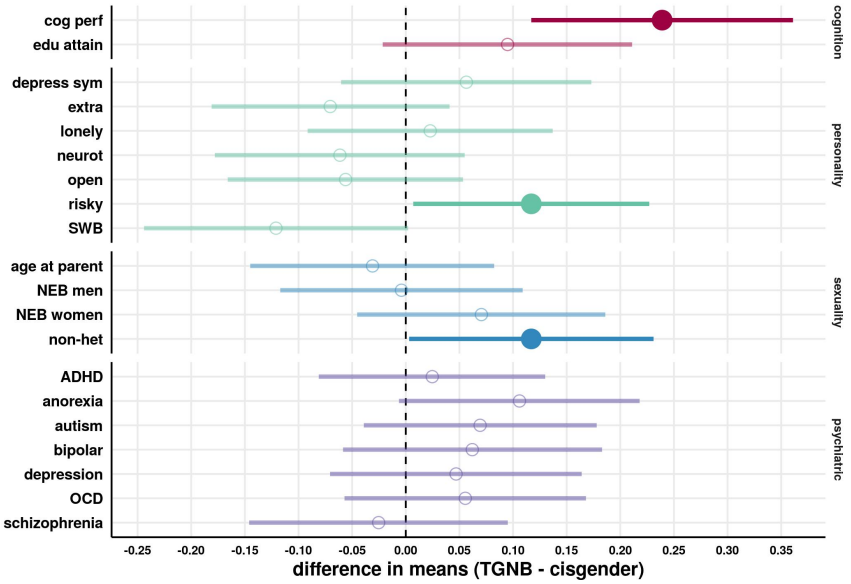

Figure 4. Replication of polygenic score differences with categorical gender identity in the larger cohort
Polygenic score differences in the larger cohort (N = 5,388). The two gender identity groups are cisgender (N = 5,089) and TGNB (N = 299). Polygenic score abbreviations: extra = extraversion, neurot = neuroticism, open = openness, risky = risky behavior, SWB = subjective well-being, NEB = number of children ever born, non-het = non-heterosexual sexual behavior.

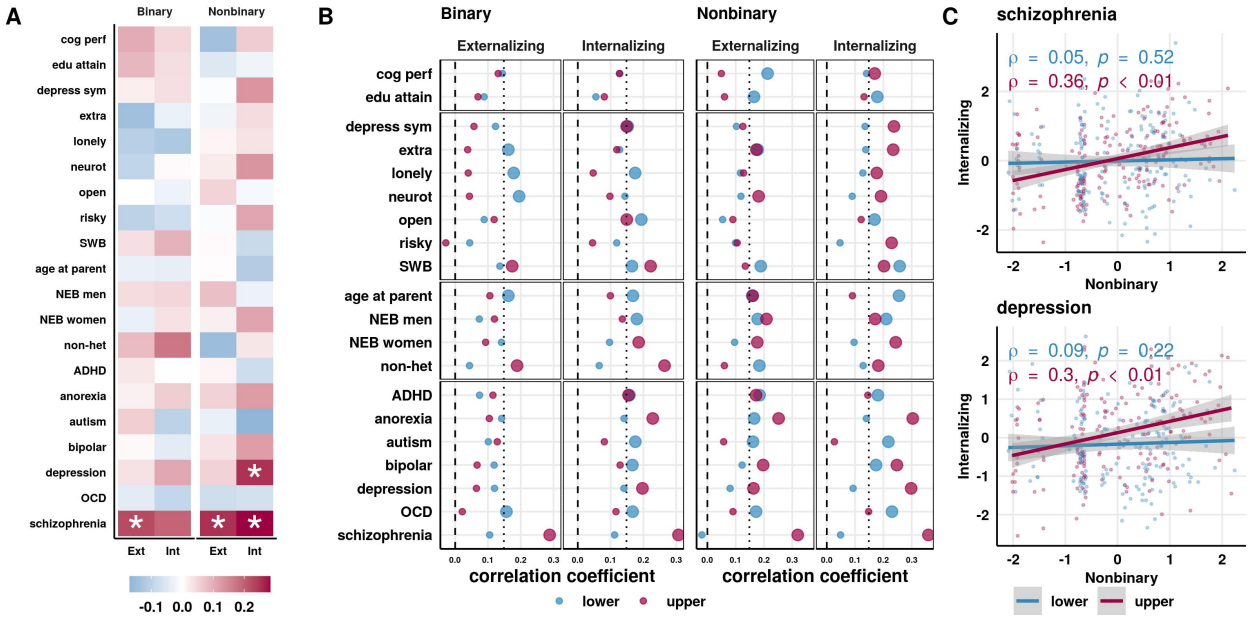

Figure 5. Correlations and interactions between gender diversity, mental health, and polygenic scores
A Polygenic score group-by-GSR interaction effects in association with the ASR values. The polygenic score groups are the upper 75th quartile (coded as 1) and the lower 25th quartile groups (coded as 0), with N = 174 in each group. The fill color is the β estimate for the interaction term from the linear model. Significant interaction terms are indicated with the white asterisk. **B** Correlations of the GSR values with the ASR values stratified by polygenic score group. The dotted lines represent the minimal correlation coefficient for statistical significance for N = 174 that is $\rho = \pm 0.148$. **C** Examples of the GSR-ASR correlations stratified by polygenic group. Polygenic score abbreviations: extra = extraversion, neurot = neuroticism, open = openness, risky = risky behavior, SWB = subjective well-being, NEB = number of children ever born, non-het = non-heterosexual sexual behavior.

Supplementary information

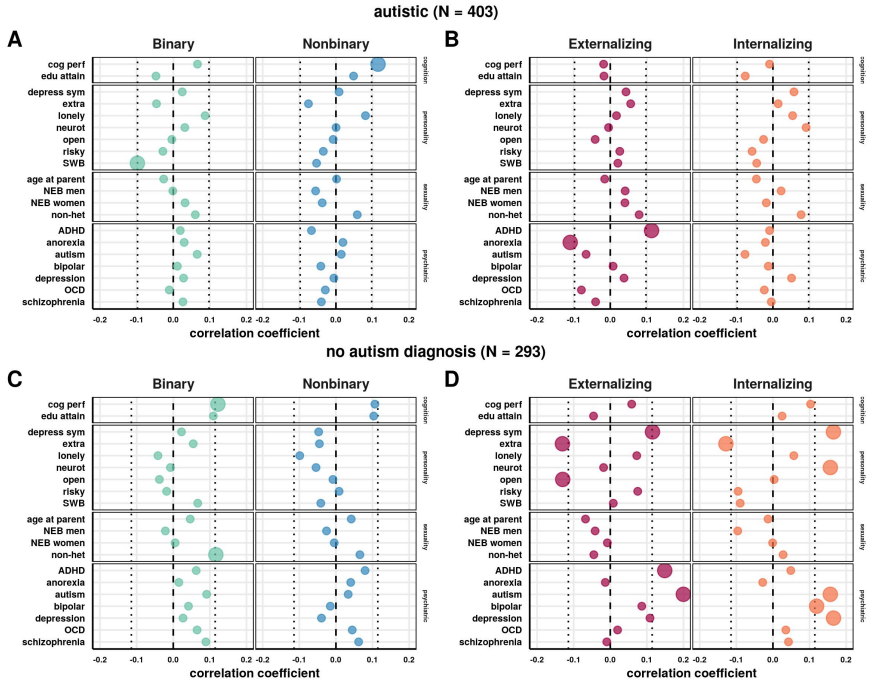

Figure S1. Polygenic score correlations with gender diversity and mental health separated by autism diagnostic status.
Polygenic score correlations in the autism subset (N = 403 out of 696) with the GSR values **A** and ASR values **B**. The dotted lines represent the minimal correlation coefficient for statistical significance that is $\rho = \pm 0.098$. Polygenic score correlations in the subset not diagnosed with autism (N = 293 out of 696) with the GSR values **C** and ASR values **D**. The dotted lines represent the minimal correlation coefficient for statistical significance that is $\rho = \pm 0.114$. Polygenic score abbreviations: extra = extraversion, neurot = neuroticism, open = openness, risky = risky behavior, SWB = subjective well-being, NEB = number of children ever born, non-het = non-heterosexual sexual behavior.

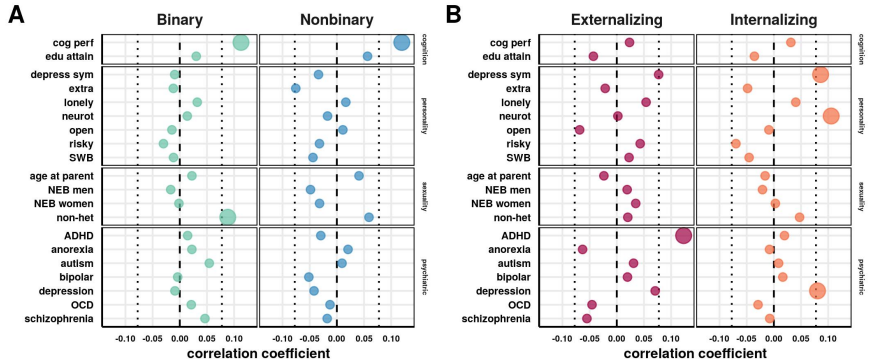

Figure S2. Polygenic score correlations with gender diversity and mental health in the Europe genetic ancestry population only (N = 639).
Polygenic score correlations with the GSR values **A** and ASR values **B** . The dotted lines represent the minimal correlation coefficient for statistical significance (nominal $p < 0.05$) that is $\rho = \pm 0.077$. Polygenic score abbreviations: extra = extraversion, neurot = neuroticism, open = openness, risky = risky behavior, SWB = subjective well-being, NEB = number of children ever born, non-het = non-heterosexual sexual behavior.

References

1. D. Westergaard, P. Moseley, F. K. H. Sørup, P. Baldi, and S. Brunak, "Population-wide analysis of differences in disease progression patterns in men and women," *Nature Communications*, vol. 10, no. 1, pp. 1–14, 2019.
2. I. Osika Friberg, G. Krantz, S. Määttä, and K. Järbrink, "Sex differences in health care consumption in Sweden: a register-based cross-sectional study," *Scandinavian journal of Public Health*, vol. 44, no. 3, pp. 264–273, 2016.
3. M. Romanelli and K. D. Hudson, "Individual and systemic barriers to health care: Perspectives of lesbian, gay, bisexual, and transgender adults.," *American Journal of Orthopsychiatry*, vol. 87, no. 6, p. 714, 2017.
4. E. A. Khramtsova, L. K. Davis, and B. E. Stranger, "The role of sex in the genomics of human complex traits," *Nature Reviews Genetics*, vol. 20, no. 3, pp. 173–190, 2019.
5. A. E. Lloyd and M. P. Galupo, "What people with normative identities believe about sex, gender and sexual orientation," *Psychology & Sexuality*, vol. 10, no. 3, pp. 269–280, 2019.
6. R. E. Morgan, C. Dragon, G. Daus, J. Holzberg, R. Kaplan, H. Menne, and M. Spiegelman, "Updates on terminology of sexual orientation and gender identity survey measures (fcsm 20-03)," *Federal Committee on Statistical Methodology*, 2020.
7. T. R. Thomas, D. Hofammann, B. G. McKenna, A. IR van der Miesen, M. A. Stokes, P. Daniolos, and J. J. Michaelson, "Community attitudes on genetic research of gender identity, sexual orientation, and mental health," *PLOS One*, vol. 15, no. 7, p. e0235608, 2020.
8. B. D. Wilson and I. H. Meyer, "Nonbinary LGBTQ adults in the United States," 2021.
9. R. A. Lippa, "Gender-related traits in gay men, lesbian women, and heterosexual men and women: The virtual identity of homosexual-heterosexual diagnosticity and gender diagnosticity," *Journal of Personality*, vol. 68, no. 5, pp. 899–926, 2000.
10. M. P. Marshal, L. J. Dietz, M. S. Friedman, R. Stall, H. A. Smith, J. McGinley, B. C. Thoma, P. J. Murray, A. R. D'Augelli, and D. A. Brent, "Suicidality and depression disparities between sexual minority and heterosexual youth: A meta-analytic review," *Journal of Adolescent Health*, vol. 49, no. 2, pp. 115–123, 2011.
11. P. B. Barr, T. B. Bigdeli, and J. L. Meyers, "Prevalence, comorbidity, and sociodemographic correlates of psychiatric disorders reported in the all of us research program," *JAMA Psychiatry*, 2022.
12. M. R. Woodford, G. Weber, Z. Nicolazzo, R. Hunt, A. Kulick, T. Coleman, S. Coulombe, and K. A. Renn, "Depression and attempted suicide among LGBTQ college students: Fostering resilience to the effects of heterosexism and cisgenderism on campus," *Journal of College Student Development*, vol. 59, no. 4, pp. 421–438, 2018.
13. A. E. Green, J. P. DeChants, M. N. Price, and C. K. Davis, "Association of gender-affirming hormone therapy with depression, thoughts of suicide, and attempted suicide among transgender and nonbinary youth," *Journal of Adolescent Health*, vol. 70, no. 4, pp. 643–649, 2022.
14. T. J. C. Polderman, B. P. C. Kreukels, M. S. Irwig, L. Beach, Y.-M. Chan, E. M. Derks, I. Esteva, J. Ehrenfeld, M. D. Heijer, D. Posthuma, L. Raynor, A. Tishelman, L. K. Davis, and International Gender Diversity Genomics Consortium, "The biological contributions to gender identity and gender diversity: Bringing data to the table," *Behavior Genetics*, vol. 48, no. 2, pp. 95–108, 2018.
15. A. Abdellaoui and K. J. H. Verweij, "Dissecting polygenic signals from genome-wide association studies on human behaviour," *Nature Human Behaviour*, vol. 5, no. 6, pp. 686–694, 2021.
16. J. G. Theisen, V. Sundaram, M. S. Filchak, L. P. Chorich, M. E. Sullivan, J. Knight, H.-G. Kim, and L. C. Layman, "The use of whole exome sequencing in a cohort of transgender individuals to identify rare genetic variants," *Scientific Reports*, vol. 9, no. 1, pp. 1–11, 2019.
17. A. R. Sanders, G. W. Beecham, S. Guo, K. Dawood, G. Rieger, R. S. Krishnappa, A. B. Kolundzija, J. M. Bailey, and E. R. Martin, "Genome-Wide Linkage and Association Study of Childhood Gender Nonconformity in Males," *Archives of Sexual Behavior*, vol. 50, pp. 3377–3383, Nov. 2021.
18. A. Ganna, K. J. H. Verweij, M. G. Nivard, R. Maier, R. Wedow, A. S. Busch, A. Abdellaoui, S. Guo, J. F. Sathirapongsasuti, 23andMe Research Team, P. Lichtenstein, S. Lundström, N. Långström, A. Auton, K. M. Harris, G. W. Beecham, E. R. Martin, A. R. Sanders, J. R. B. Perry, B. M. Neale, and B. P. Zietsch, "Large-scale GWAS reveals insights into the genetic architecture of same-sex sexual behavior," *Science*, vol. 365, no. 6456, 2019.
19. A. G. Cuevas, F. D. Mann, D. R. Williams, and R. F. Krueger, "Discrimination and anxiety: Using multiple polygenic scores to control for genetic liability," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, no. 1, 2021.
20. P. Feliciano, A. M. Daniels, L. G. Snyder, A. Beaumont, A. Camba, A. Esler, A. G. Gulsrud, A. Mason, A. Gutierrez, A. Nicholson, *et al.*, "Spark: a us cohort of 50,000 families to accelerate autism research," *Neuron*, vol. 97, no. 3, pp. 488–493, 2018.
21. E. Hisle-Gorman, C. A. Landis, A. Susi, N. A. Schvey, G. H. Gorman, C. M. Nylund, and D. A. Klein, "Gender dysphoria in children with autism spectrum disorder," *LGBT Health*, vol. 6, no. 3, pp. 95–100, 2019.
22. V. Warrier, D. M. Greenberg, E. Weir, C. Buckingham, P. Smith, M.-C. Lai, C. Allison, and S. Baron-Cohen, "Elevated rates of autism, other neurodevelopmental and psychiatric diagnoses, and autistic traits in transgender and gender-diverse individuals," *Nature Communications*, vol. 11, no. 1, pp. 1–12, 2020.
23. T. M. Achenbach and L. Rescorla, "Manual for the ASEBA adult forms & profiles," 2003.
24. . G. P. Consortium *et al.*, "A global reference for human genetic variation," *Nature*, vol. 526, no. 7571, p. 68, 2015.
25. J. F. Strang, G. L. Wallace, G. Slesaransky-Poe, and J. S. Yang, "The gender self-report (gsr)," 2021.
26. J. J. Lee, R. Wedow, A. Okbay, E. Kong, O. Maghzian, M. Zacher, T. A. Nguyen-Viet, P. Bowers, J. Sidorenko, R. K. Linnér, *et al.*, "Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals," *Nature Genetics*, vol. 50, no. 8, pp. 1112–1121, 2018.
27. S. M. van den Berg, M. H. de Moor, K. J. Verweij, R. F. Krueger, M. Luciano, A. A. Vasquez, L. K. Matteson, J. Derringer, N. Amin, S. D. Gordon, *et al.*, "Meta-analysis of genome-wide association studies for extraversion: Findings from the genetics of personality consortium," *Behavior Genetics*, vol. 46, no. 2, pp. 170–182, 2016.

28. M. Nagel, P. R. Jansen, S. Stringer, K. Watanabe, C. A. De Leeuw, J. Bryois, J. E. Savage, A. R. Hammerschlag, N. G. Skene, A. B. Muñoz-Manchado, *et al.*, "Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways," *Nature Genetics*, vol. 50, no. 7, pp. 920–927, 2018.

29. M. H. De Moor, P. T. Costa, A. Terracciano, R. F. Krueger, E. J. De Geus, T. Toshiko, B. W. Penninx, T. Esko, P. A. Madden, J. Derringer, *et al.*, "Meta-analysis of genome-wide association studies for personality," *Molecular Psychiatry*, vol. 17, no. 3, pp. 337–349, 2012.

30. D. M. Howard, M. J. Adams, T.-K. Clarke, J. D. Hafferty, J. Gibson, M. Shirali, J. R. Coleman, S. P. Hagenaars, J. Ward, E. M. Wigmore, *et al.*, "Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions," *Nature Neuroscience*, vol. 22, no. 3, pp. 343–352, 2019.

31. Neale Lab, "Uk biobank gwas - round 2," 2018.

32. R. K. Linnér, P. Biroli, E. Kong, S. F. W. Meddens, R. Wedow, M. A. Fontana, M. Lebreton, S. P. Tino, A. Abdellaoui, A. R. Hammerschlag, *et al.*, "Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences," *Nature Genetics*, vol. 51, no. 2, pp. 245–257, 2019.

33. A. Okbay, B. M. Baselmans, J.-E. De Neve, P. Turley, M. G. Nivard, M. A. Fontana, S. F. W. Meddens, R. K. Linnér, C. A. Rietveld, J. Derringer, *et al.*, "Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses," *Nature Genetics*, vol. 48, no. 6, pp. 624–633, 2016.

34. N. Barban, R. Jansen, R. De Vlaming, A. Vaez, J. J. Mandemakers, F. C. Tropf, X. Shen, J. F. Wilson, D. I. Chasman, I. M. Nolte, *et al.*, "Genome-wide analysis identifies 12 loci influencing human reproductive behavior," *Nature Genetics*, vol. 48, no. 12, pp. 1462–1472, 2016.

35. D. Demontis, R. K. Walters, J. Martin, M. Mattheisen, T. D. Als, E. Agerbo, G. Baldursson, R. Belliveau, J. Bybjerg-Grauholm, M. Bækvad-Hansen, *et al.*, "Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder," *Nature Genetics*, vol. 51, no. 1, pp. 63–75, 2019.

36. H. J. Watson, Z. Yilmaz, L. M. Thornton, C. Hübel, J. R. Coleman, H. A. Gaspar, J. Bryois, A. Hinney, V. M. Leppä, M. Mattheisen, *et al.*, "Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa," *Nature Genetics*, vol. 51, no. 8, pp. 1207–1214, 2019.

37. J. Grove, S. Ripke, T. D. Als, M. Mattheisen, R. K. Walters, H. Won, J. Pallesen, E. Agerbo, O. A. Andreassen, R. Anney, *et al.*, "Identification of common genetic risk variants for autism spectrum disorder," *Nature Genetics*, vol. 51, no. 3, pp. 431–444, 2019.

38. N. Mullins, A. J. Forstner, K. S. O'Connell, B. Coombes, J. R. Coleman, Z. Qiao, T. D. Als, T. B. Bigdeli, S. Børte, J. Bryois, *et al.*, "Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology," *Nature Genetics*, vol. 53, no. 6, pp. 817–829, 2021.

39. D. M. Howard, M. J. Adams, T.-K. Clarke, J. D. Hafferty, J. Gibson, M. Shirali, J. R. Coleman, S. P. Hagenaars, J. Ward, E. M. Wigmore, *et al.*, "Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions," *Nature Neuroscience*, vol. 22, no. 3, pp. 343–352, 2019.

40. P. D. Arnold, K. D. Askland, C. Barlassina, L. Bellodi, O. Bienvenu, D. Black, M. Bloch, H. Brentani, C. L. Burton, B. Camarena, *et al.*, "Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis," *Molecular psychiatry*, vol. 23, no. 5, pp. 1181–1181, 2018.

41. S. Ripke, J. T. Walters, M. C. O'Donovan, S. W. G. of the Psychiatric Genomics Consortium, *et al.*, "Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia," *MedRxiv*, 2020.

42. J. D. Bosse and L. Chiodo, "It is complicated: gender and sexual orientation identity in LGBTQ youth," *Journal of Clinical Nursing*, vol. 25, no. 23-24, pp. 3665–3675, 2016.

43. J. F. Strang, A. I. van der Miesen, R. Caplan, C. Hughes, S. daVanport, and M.-C. Lai, "Both sex-and gender-related factors should be considered in autism research and clinical practice," 2020.

44. G. Schwartz, R. M. Kim, A. B. Kolundzija, G. Rieger, and A. R. Sanders, "Biodemographic and physical correlates of sexual orientation in men," *Archives of sexual behavior*, vol. 39, no. 1, pp. 93–109, 2010.

45. B. P. Zietsch, M. J. Sidari, A. Abdellaoui, R. Maier, N. Långström, S. Guo, G. W. Beecham, E. R. Martin, A. R. Sanders, and K. J. H. Verweij, "Genomic evidence consistent with antagonistic pleiotropy may help explain the evolutionary maintenance of same-sex sexual behaviour in humans," *Nature Human Behaviour*, vol. 5, no. 9, pp. 1251–1258, 2021.

46. Y. Fang, L. Scott, P. Song, M. Burmeister, and S. Sen, "Genomic prediction of depression risk and resilience under stress," *Nature Human Behaviour*, vol. 4, no. 1, pp. 111–118, 2019.

47. T.-K. Clarke, M. Lupton, A. M. Fernandez-Pujals, J. Starr, G. Davies, S. Cox, A. Pattie, D. Liewald, L. Hall, D. MacIntyre, *et al.*, "Common polygenic risk for autism spectrum disorder (ASD) is associated with cognitive ability in the general population," *Molecular Psychiatry*, vol. 21, no. 3, pp. 419–425, 2016.

48. S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. de Bakker, M. J. Daly, and P. C. Sham, "PLINK: A tool set for whole-genome association and population-based linkage analyses," *The American Journal of Human Genetics*, vol. 81, no. 3, pp. 559–575, 2007.

49. D. Taliun, D. N. Harris, M. D. Kessler, J. Carlson, Z. A. Szpiech, R. Torres, S. A. G. Taliun, A. Corvelo, S. M. Gogarten, H. M. Kang, *et al.*, "Sequencing of 53,831 diverse genomes from the NHLBI TOPMed program," *Nature*, vol. 590, no. 7845, pp. 290–299, 2021.

50. S. Das, L. Forer, S. Schönherr, C. Sidore, A. E. Locke, A. Kwong, S. I. Vrieze, E. Y. Chew, S. Levy, M. McGue, *et al.*, "Next-generation genotype imputation service and methods," *Nature Genetics*, vol. 48, no. 10, pp. 1284–1287, 2016.

51. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, *et al.*, "Twelve years of SAMtools and BCFtools," *Gigascience*, vol. 10, no. 2, p. giab008, 2021.

52. F. Privé, H. Aschard, A. Ziyatdinov, and M. G. Blum, "Efficient analysis of large-scale genome-wide data with two r packages: bigstatsr and bigsnpr," *Bioinformatics*, vol. 34, no. 16, pp. 2781–2787, 2018.

53. F. Privé, K. Luu, M. G. B. Blum, J. J. McGrath, and B. J. Vilhjálmsson, "Efficient toolkit implementing best practices for principal component analysis of population genetic data," *Bioinformatics*, vol. 36, no. 16, pp. 4449–4457, 2020.

54. A. C. F. Lewis, S. J. Molina, P. S. Appelbaum, B. Dauda, A. D. Rienzo, A. Fuentes, S. M. Fullerton, N. A. Garrison, N. Ghosh, E. M. Hammonds, D. S. Jones, E. E. Kenny, P. Kraft, S. S.-J. Lee, M. Mauro, J. Novembre, A. Panofsky, M. Sohail, B. M. Neale, and D. S. Allen, "Getting genetic ancestry right for science and society," *Science*, vol. 376, no. 6590, pp. 250–252, 2022.

55. J. Yang, S. H. Lee, M. E. Goddard, and P. M. Visscher, "GCTA: A tool for genome-wide complex trait analysis," *The American Journal of Human Genetics*, vol. 88, no. 1, pp. 76–82, 2011.

56. F. Privé, J. Arbel, and B. J. Vilhjálmsson, "Ldpred2: better, faster, stronger," *Bioinformatics*, vol. 36, no. 22-23, pp. 5424–5431, 2020.

57. R Core Team, *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2013.

58. J. Becker, C. A. Burik, G. Goldman, N. Wang, H. Jayashankar, M. Bennett, D. W. Belsky, R. K. Linnér, R. Ahlskog, A. Kleinman, *et al.*, "Resource profile and user guide of the polygenic index repository," *Nature Human Behaviour*, pp. 1–15, 2021.

59. S. Champely, *pwr: Basic Functions for Power Analysis*, 2020. R package version 1.3-0.

60. American Psychiatric Association, "A guide for working with transgender and gender nonconforming patients," 2017.

Polygenic scores clarify the relationship between mental health and gender diversity

Taylor R. Thomas[1], Ashton J. Tener[1], Amy M. Pearlman[2], Katherine L. Imborek[3], Ji Seung Yang[4], John F. Strang[5], Jacob J. Michaelson[1,6,7*]

Public summary

The way we act (behavior) is influenced by how our brains grow and function. Some of the ways our brains grow and function are influenced by our genes (DNA). Everyone has slightly different versions of DNA. This is a normal part of human diversity. Some of these DNA differences lead to differences in our brains. Brain differences can lead to behavior differences like in personality, intelligence, or mental health.

In this study, we asked whether our DNA is involved in gender identity and gender expression. We use the term "gender diversity" to mean differences in gender identity or gender expression. People with greater gender diversity are more likely to be transgender and/or nonbinary, although cisgender people can also have differences in gender expression. People with greater gender diversity are also more likely to be autistic, so we conducted this study with the help of SPARK participants (SPARK is the largest study of autism). Approximately half of our study participants were autistic adults, and the others were not autistic but do have an immediate relative who is autistic.

What we found

We found that thousands of DNA differences, when combined, are linked to differences in gender diversity. Specifically, we found that DNA differences linked to higher intelligence were also linked to greater gender diversity. We need to do more research to understand why DNA differences linked to higher intelligence are also linked to greater gender diversity.

Gender diverse people are at increased risk for stress due to discrimination. Scientists have repeatedly found that gender diverse people are at a greater risk for poorer mental health. In this study, we showed that greater gender diversity is linked to poorer mental health only among people who have high genetic risk for psychiatric conditions. Greater gender diversity was not linked to poorer mental health among those with low genetic risk for psychiatric conditions. This may mean that their DNA differences help them to be resilient against minority stress.

What our study does not show

We did not identify, nor attempt to identify, a "transgender" or "nonbinary" gene. We cannot predict a person's gender diversity from their DNA. We found very little evidence that the DNA differences linked to major psychiatric conditions are also linked to gender diversity. Larger studies in the future may be able to identify weaker effects, but our study does not support a strong genetic connection between psychiatric conditions and gender diversity. Gender diversity is not purely genetic, but genetic factors do play a role.

Why this study is important

This is one of the first genetic studies of gender diversity. Many people say that gender is a purely social construct, with no biological factors involved. Our results show that the DNA differences linked to higher intelligence are also linked to greater gender diversity. We also show that whether minority stress translates into poorer mental health depends on the person's level of genetic risk for psychiatric conditions. Ultimately, we believe this line of research will advance the health of gender diverse people through a greater understanding of how genetics interact with gender diversity in determining health outcomes.