Corresponding author(s): Andrew Grotzinger
e-mail: agrotzin@utexas.edu

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒ Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	Lead SNPs and GWAS for simulation studies were conducted using PLINK v1.9. Simulation of phenotypes with corresponding genotypes was conducting using GCTA v1.91. The construction of polygenic scores for out-of-sample prediction was conducted using PRSice v1.25. Genomic SEM was run using code we make available for download at https://github.com/MichelNivard/GenomicSEM

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All summary statistics used are available for public download. No raw data was used except for the construction of polygenic scores (PGSs) for out-of-sample prediction in UK Biobank and Generation Scotland. Although we are unable to make this data available ourselves, the data may be requested by others. Supplementary Table 1 provides references for all of the summary statistics used in our analyses. All summary statistics used are available for public use, and links are provide directly below.

For construction of the p-factor, summary statistics for PTSD, Anxiety, Bipolar Disorder, and Major Depressive Disorder were downloaded from: http://www.med.unc.edu/pgc/results-and-downloads. Summary statistics for Schizophrenia were downloaded from: http://walters.psycm.cf.ac.uk/.

The GWIS summary statistics for Bipolar Disorder and Schizophrenia were also downloaded from the cross-disorder section of the PGC website. The summary statistics for educational achievement are available for download from: https://www.thessgac.org/data.

Summary statistics for item-level indicators of neuroticism were downloaded from: https://docs.google.com/spreadsheets/d/1b3oGI2lUt57BcuHttWaZotQcI0-mBRPyZihz87Ms_No/edit#gid=1209628142. We use Round 1 of the Neale's Lab UKB GWAS results.

Summary statistics for early-life traits used in the factor model of anthropometric traits (birth length, birth weight, infant head circumference, childhood obesity) were obtained from: https://egg-consortium.org/.

Summary statistics used for the anthropometric traits example for BMI, waist-hip ratio, hip circumference, waist circumference, and height are available from: https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Our study uses publically available summary statistics to examine a new developed method, Genomic SEM. Genomic SEM can be used to examine any form of structural equation model (e.g., factor analysis, mediation, etc.) using summary statistics. The software provides model fit statistics, and can also be used to produce summary statistics for a latent trait.
Research sample	We use only publically available summary statistics for our analyses. As Genomic SEM relies on ld-score regression (LDSC) to construct genetic covariance matrices, and LDSC requires summary statistics to be within a single ethnic population due to differences in linkage disequilibrium across populations, we use only summary statistics restricted to European populations.
Sampling strategy	This is not applicable as we use previously collected data.
Data collection	This is not applicable as we use previously collected data.
Timing	This is not applicable as we use previously collected data.
Data exclusions	We use only summary statistics from European populations due to the requirements of LDSC, as noted above.
Non-participation	This is not applicable as we use previously collected data.
Randomization	This is not applicable as we use previously collected data.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging