

The pleiotropic architecture of human impulsivity across biological scales

Travis T. Mallard^{1,2,3}, Justin D. Tubbs^{1,2,3}, Mariela Jennings⁴, Yingzhe Zhang⁵, Daniel E. Gustavson⁶, Andrew D. Grotzinger^{6,7}, Margaret L. Westwater⁸, Camille M. Williams⁹, Rebecca G. Fortgang^{2,10}, 23andMe Research Team, Sarah L. Elson¹¹, Pierre Fontanillas¹¹, Lea K. Davis^{12,13,14}, Armin Raznahan¹⁵, Elliot M. Tucker-Drob^{9,16}, Karmel W. Choi^{1,2,3}, Tian Ge^{1,2,3}, Jordan W. Smoller^{1,2,3,10}, Abraham A. Palmer^{4,17}, & Sandra Sanchez-Roige^{4,12}

Affiliations:

1. Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA
2. Department of Psychiatry, Harvard Medical School, Boston, MA
3. Broad Institute of MIT and Harvard, Boston, MA
4. Department of Psychiatry, University of California San Diego, La Jolla, CA
5. Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA
6. Institute for Behavioral Genetics, University of Colorado Boulder, Boulder, CO
7. Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO
8. Department of Psychiatry, University of Oxford, Oxford, UK
9. Population Research Center, University of Texas at Austin, Austin, TX
10. Center for Precision Psychiatry, Massachusetts General Hospital, Boston, MA
11. 23andMe Inc., Sunnyvale, CA
12. Department of Medicine, Division of Genetic Medicine, Vanderbilt University Medical Center, Nashville, TN
13. Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center, Nashville, TN
14. Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN
15. Section on Developmental Neurogenomics, National Institute of Mental Health, Bethesda, MD
16. Department of Psychology, University of Texas at Austin, Austin, TX
17. Institute for Genomic Medicine, University of California San Diego, La Jolla, CA

Main Text

Word Count: 5,024 in Main Text; 3,187 in Methods
Tables: 1
Figures: 5

Corresponding Author

Name: Travis T. Mallard, Ph.D.
Email Address: tmallard@mgh.harvard.edu
Mailing Address: Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, 185 Cambridge Street, Boston, MA 02114

Abstract

Impulsivity is a complex psychological construct that represents a core feature of many psychiatric and neurological conditions. Here, we used multivariate methods to formally model the genetic architecture of impulsivity in humans, advancing genomic discovery and revealing pervasive pleiotropy that largely counters theories of impulsivity as a unitary construct. We identified 18 loci and 93 genes with diverse effects in GWAS and TWAS analyses, respectively, including a hotspot at 17q21.31 that harbors genes involved in neurodevelopmental and neurodegenerative disorders. Downstream analyses revealed that heterogeneous signals were localized to specific biological correlates, including expression in brain tissue during fetal development and cortical alterations in the inferior frontal gyrus. Polygenic score analyses suggested that liability for different forms of impulsivity may differentiate across development, operating via broad pathways early in life but affecting diverse outcomes by adulthood. Collectively, our study generates new insights into the pleiotropic architecture of impulsivity, which provides a more comprehensive understanding of its multi-faceted biology.

Introduction

Impulsivity is a complex psychological construct that has a considerable role in health and behavior. Broadly characterized as a predisposition towards rapid, unplanned responses to stimuli¹, impulsivity has been implicated in multiple psychiatric conditions^{2,3}, maladaptive decision-making^{4,5}, and risk-taking behaviors⁶. With well-established links to brain structure and function^{4,7}, it is considered to be an endophenotype that mediates genetic effects on diverse cognitive, emotional, and behavioral domains across various clinical and non-clinical populations^{3,8-12}. Given its cross-cutting relevance, a deeper etiological understanding of impulsivity could improve patient risk stratification and facilitate the development of novel therapeutic approaches with transdiagnostic applications¹³⁻¹⁵. However, there remains considerable debate surrounding its definition, dimensionality, and optimal measurement^{1,16}, which presents a challenge to achieving these aims.

Insights from neuroscience research increasingly suggest that impulsivity can be ‘split’ into related-but-distinct facets rather than considered a unitary trait. These different aspects, such as motor and choice impulsivity, have been found to reflect only partially shared neural circuitry⁷. This recognition has spurred the development of numerous theoretical models, including reinforcement sensitivity theory¹⁷⁻¹⁹ and the dual systems model²⁰⁻²², that attempt to explain the heterogeneous biology underlying impulsivity in humans. Although considerable progress has been made in this domain, the etiology of impulsivity remains insufficiently characterized. Disentangling shared and unique biological influences on different manifestations of impulsivity will require comprehensive multi-modal data and integrative approaches^{13,23}.

In contrast to these neuroscientific findings, advances in the fields of clinical psychology, psychiatry, and genetics have revealed that many complex mental constructs can be effectively ‘lumped’ into broader categories at multiple levels of analysis. For example, multivariate models have provided compelling evidence for a hierarchical structure of mental disorders, where higher-order dimensions of psychopathology, such as thought disorder, internalizing, and externalizing, explain substantial phenotypic and genetic covariance in psychiatric diagnoses or symptoms²⁴⁻²⁷. This perspective has transformed our understanding of psychiatric conditions, leading to the identification of shared genetic influences, prioritizing novel transdiagnostic therapeutic targets, and accelerating genomic discovery for understudied conditions²⁸⁻³³. Considering the success of this approach in explaining the widespread pleiotropy observed among similarly complex traits³⁴⁻³⁶, it stands to reason that genetic studies of impulsivity may similarly benefit from lumping³⁷.

With the recent completion of a large-scale genome-wide association study (GWAS) of multiple measures of impulsivity (maximum $N = 133,517$)³⁸, we now have sufficient data to characterize the joint genetic architecture of multiple forms of impulsivity. In the present study, we conducted a multivariate GWAS of eight facets measured with two psychological instruments: the Barratt Impulsiveness Scale³⁹ and the UPPS-P Impulsive Behavior Scale⁴⁰ (see **Method** for further description of phenotypes). Specifically, we use the Genomic Structural Equation Modeling (Genomic SEM)²⁹ framework to test the ‘splitting’ versus ‘lumping’ hypotheses with two distinct statistical models: (1) an “omnibus” model that provides an unstructured meta-analytic test of association, and (2) a common factor of impulsivity, in which associations operate through a latent genetic factor. We evaluate

these models and their outputs across molecular genetic, transcriptomic, neurogenomic, and phenomic levels to gain a more comprehensive understanding of the underpinnings and correlates of impulsivity across biological scales. By doing so, we discover new insights into the complex etiology of impulsivity that advance understanding of its multifaceted nature and relationship with health and well-being.

Results

Genomic factor analysis reveals widespread pleiotropy among facets of impulsivity

To characterize the genomic relationships between impulsivity facets (see **Table 1** for overview), we first used linkage disequilibrium (LD) score regression to estimate the genetic covariances among study phenotypes. Generally, BIS and UPPS-P phenotypes exhibited patterns of moderate-to-large genetic overlap (mean $r_g = 0.434$; **Figure 1a**), except for sensation seeking which had a more modest overlap with other facets (maximum $|r_g| = 0.337$). Examination of the standardized genetic covariance matrix revealed that the first eigenvector explained 54% of the total genetic variance among these traits (**Figure 1b**). Consistent with this observation, we found that a common factor model provided a good fit to the data ($\chi^2[17] = 325.079$, CFI = 0.947, SRMR = 0.08; **Figure S1**; see **Supplementary Note** for additional details on the factor analysis). Accordingly, we proceeded with the omnibus and common factor models, as described below (**Figure 1c,d**).

Table 1. Summary of eight impulsivity GWAS included in the present study.

Impulsivity Facet	N	h^2	λ_{GC}	Mean χ^2	Intercept	Ratio	# of loci
Attentional (ATTN)	124,739	0.070	1.147	1.201	1.026	0.127	2
Motor (MOTR)	124,104	0.067	1.147	1.178	1.011	0.064	3
Negative Urgency (NEGU)	132,559	0.079	1.200	1.236	1.017	0.071	5
Nonplanning (NONP)	123,509	0.093	1.200	1.252	1.022	0.087	6
(Lack of) Perseverance (PERS)	133,517	0.063	1.147	1.186	1.018	0.094	2
Positive Urgency (POSU)	132,132	0.061	1.147	1.182	1.017	0.093	2
(Lack of) Premeditation (PREM)	132,667	0.065	1.147	1.175	1.003	0.017	2
Sensation Seeking (SENS)	132,395	0.097	1.200	1.265	0.998	< 0	7

Note: All GWAS summary statistics originate from Sanchez-Roige and colleagues' univariate study of impulsivity facets³⁸. The statistics in this table were calculated with linkage disequilibrium score regression except for the number of loci, which were defined using the same pipeline as the multivariate results (**Methods**). N = GWAS sample size; h^2 = observed SNP heritability; λ_{GC} = genomic inflation factor; mean χ^2 = average χ^2 statistic; intercept = estimated intercept from linkage disequilibrium score regression; ratio = the attenuation ratio, which is defined as $(\text{intercept} - 1)/(\text{mean } \chi^2 - 1)$; # of loci = number of independent genome-wide significant loci.

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

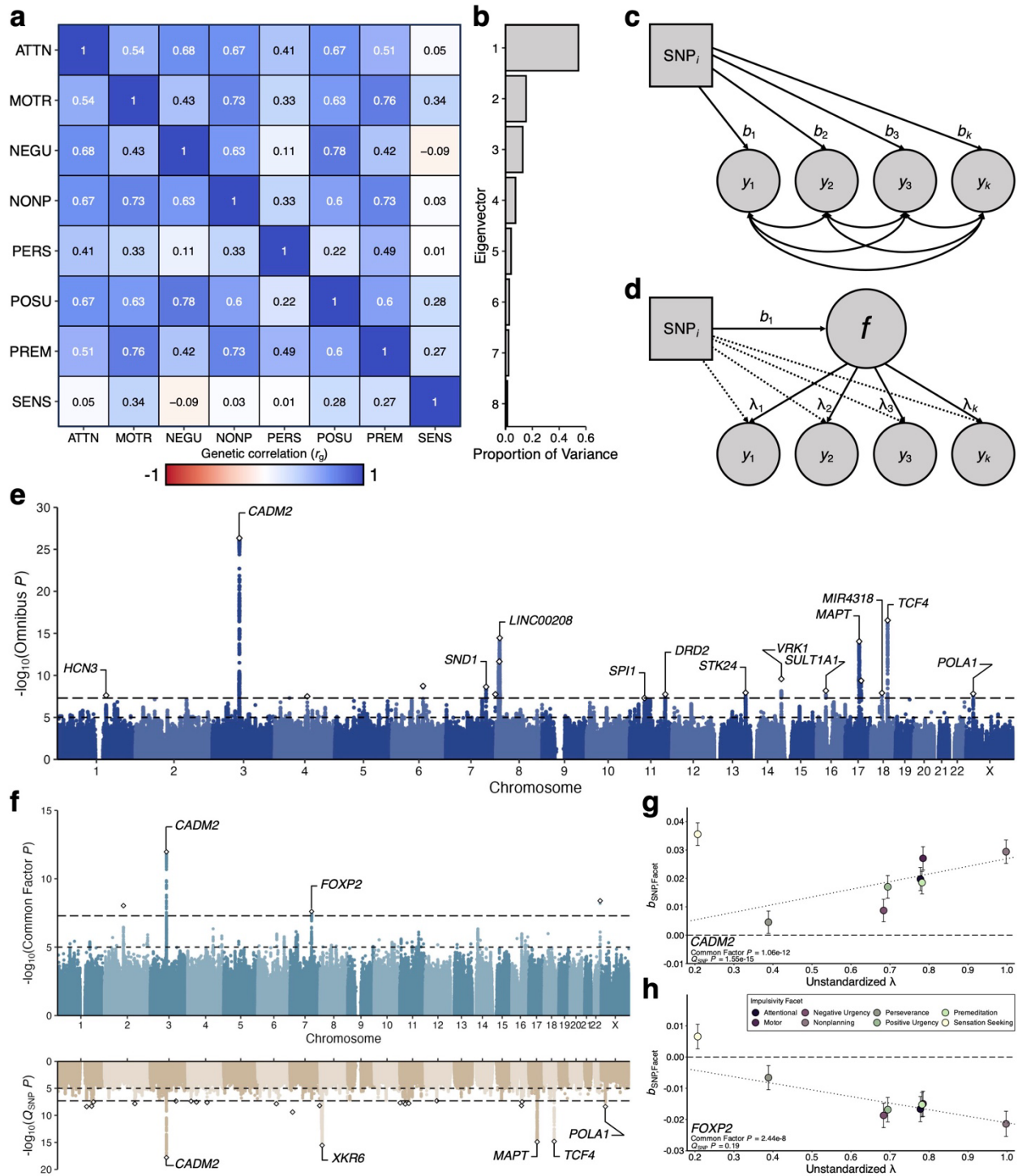


Figure 1. Modeling the multivariate genetic architecture of impulsivity. **a**, Matrix of genetic correlations (r_g) among the eight impulsivity facets, as estimated with LD score regression. **b**, Bar chart illustrating the proportion of variance explained by each eigenvector from the genetic correlation matrix. **c,d**, Path diagrams of the two discovery models employed in the present study. The omnibus model (**c**) provides an unstructured multivariate test of association across study phenotypes ($y_1 - y_k$), while the common factor model (**d**) estimates SNP effects that operate through latent genetic factor f . Circles represent the inferred genetic components of model indicators and the common factor. One-headed arrows pointing from independent variables to dependent variables reflect regression relationships (regression coefficients as $b_1 - b_k$, factor loadings as $\lambda_1 - \lambda_k$). Dotted one-headed arrows refer to the independent pathways examined in the Q_{SNP} test. Two-headed arrows connecting variables represent covariance relationships. For simplicity, (residual) variables are not depicted. **e,f**, Manhattan (**e**) and Miami (**f**) plots of GWAS results for the omnibus and common factor models, respectively, where the x-axis refers to chromosomal position, the y-axis refers to statistical significance on a $-\log_{10}$ scale, the horizontal dashed line denotes genome-wide significance ($P = 5e-8$), and the horizontal dotted line marks suggestive significance ($P = 1e-5$). Lead SNPs of independent genomic risk loci are represented by white diamonds. **g,h**, Scatter plots depicting the indicator GWAS estimates and corresponding standard errors for the lead *CADM2* (**g**) and *FOXP2* (**h**) variants as a function of unstandardized factor loadings.

Multivariate GWAS and TWAS identifies novel risk genes with significant heterogeneity

Using Genomic SEM, we conducted a series of multivariate GWAS ($n_{\text{SNPs}} = 7,053,212$) and transcriptome-wide association study (TWAS; $n_{\text{gene-tissue pairs}} = 51,698$) analyses to investigate the joint genomic architecture of impulsivity in a sample of 134,935 individuals. Specifically, we employed two models for gene discovery: (1) an ‘omnibus’ model (i.e., an unstructured meta-analytic test of association) (**Figure 1c**), and (2) the common factor of impulsivity described above, where SNP- and gene-level associations operate through a latent genetic factor (**Figure 1d**). In other words, for each SNP or gene, the omnibus model tested whether the overall pattern of effects was non-zero, while the common factor model tested whether the SNP or gene influenced the indicators via a common pathway (the latent factor). Q_{SNP} and Q_{Gene} tests were also used to evaluate the degree to which effects were (in)consistent with the common pathway, providing an index of heterogeneity at each level of analysis (see **Methods** for further description and comparison of all models).

We observed substantial inflation of the GWAS test statistics for both the omnibus ($\lambda_{\text{GC}} = 1.592$, mean $\chi^2 = 1.535$) and common factor ($\lambda_{\text{GC}} = 1.22$, mean $\chi^2 = 1.237$) models (**Figure 1e,f**). The LD score regression intercepts and attenuation ratios indicated that test-statistic inflation was primarily due to robust polygenic architecture as opposed to population stratification or other confounding (omnibus: intercept = 1.021 [s.e. = 0.009], attenuation ratio = 0.039 [s.e. = 0.017]; common factor: intercept = 1.008 [s.e. = 0.008], attenuation ratio = 0.033 [s.e. = 0.034]). However, the substantially larger mean χ^2 of the omnibus results suggests that there is an appreciable signal that is not captured by the common factor. This inference is further supported by the Q_{SNP} statistics ($\lambda_{\text{GC}} = 1.607$, mean $\chi^2 = 1.568$), which also show a stronger signal than the common factor GWAS (**Figure S2**).

With respect to the omnibus GWAS results, we identified 18 independent ($r^2 < 0.1$) genome-wide significant loci after applying a standard clumping algorithm (**Figure 1e, Table S1**), 11 of which were novel associations that were not observed in any of the facet-level results. Some of these novel risk loci include highly pleiotropic genes, such as: a locus indexed by intronic variant rs4411372 ($P_{\text{omnibus}} = 1.16\text{e-}8$) near the *STK24* gene on chromosome 13, which has been broadly linked to cognitive and psychiatric outcomes⁴¹⁻⁴³; a locus tagged by intergenic variant rs6507215 on chromosome 14 near *VRK1*, a gene that has been previously associated with alcohol consumption and insomnia^{43,44}; and another indexed by rs6507215 ($P_{\text{omnibus}} = 1.25\text{e-}8$), an intergenic variant near *MIR4318*, which has been linked to a variety of internalizing phenotypes^{45,46}. All remaining loci are reported in **Table S1**.

We found markedly fewer associated loci in the common factor GWAS. After applying a standard clumping algorithm, only four independent genomic loci were associated with a general liability toward impulsivity (**Figure 1f, Table S2**). Two signals were particularly robust: (1) rs1865250 ($b = 0.028$, s.e. = 0.004, $P = 1.06\text{e-}12$), an intronic variant in *CADM2* that was associated with several of the constituent impulsivity phenotypes and other related behaviors³⁸; and (2) rs1563408 ($b = -0.021$, s.e. = 0.004, $P = 2.44\text{e-}8$), an intronic variant in *FOXP2*, a gene that has been previously linked to neurodevelopmental and externalizing conditions^{31,47-49}, but not individual differences in impulsivity *per se*. Although subsequent Q_{SNP} tests highlighted a significant

degree of heterogeneity in the *CADM2* locus (**Figure 1f,g**), the effects of the *FOXP2* variant on impulsivity facets were consistent with the common pathway model (**Figure 1f,h**). The other two genome-wide significant loci on chromosomes 2 and 22 consisted of only two SNPs, which is not indicative of a robust association (**Table S2**). Overall, these results provided limited genome-wide evidence for a common factor of impulsivity, which was further corroborated by the greater number of Q_{SNP} loci relative to GWAS loci (**Table S3**), as well as the finding that most Q_{SNP} signals were driven by multiple facets (**Figure S3**).

When examining effects at the level of gene expression via multivariate TWAS, we found a similar pattern of results. In the omnibus TWAS, we identified 325 gene-tissue pairs (88 unique genes) that were significantly associated with impulsivity (**Figure 2a, Table S4**). In contrast, no genes were significantly associated with a general liability to impulsivity in the common factor TWAS (**Figure 2b, Table S4**). Consistent with the stark contrast between our discovery models, Q_{Gene} analyses revealed that 96% (311/325 gene-tissue pairs) of the associations identified in the omnibus TWAS had significantly heterogeneous effects across the impulsivity facets (**Table S4**).

Given the strong concordance of effects across brain tissues (all cross-tissue $r_s > 0.856$), we used an aggregated Cauchy association test (ACAT) to integrate associations across all available tissues and further improve statistical power (**Methods**). The results of these multi-tissue TWAS analyses identified a total of 93, 1, and 92 unique genes associated with impulsivity in the omnibus, common factor, and Q_{Gene} models, respectively (**Table S5**). In the omnibus results, some of the strongest associations were concentrated within chromosome 17 at cytogenetic region 17q21.31, with *MAPT* and its non-coding RNA transcripts (*MAPT-AS1* and *MAPT-IT1*), *KANSL1* and its antisense transcript (*KANSL1-AS1*), and *CRHRI* all showing robust links to impulsivity (all ACAT $P_{\text{omnibus}} < 7.79\text{e-}14$; **Figure 2c**). Notably, these genes are located in a genetically complex segment of 17q21.31 harboring a common inversion polymorphism^{50,51} that has been linked to neurodevelopmental and neurodegenerative disorders⁵²⁻⁵⁶, as well as individual differences in brain structure⁵⁷. For the common factor of impulsivity, the only transcriptome-wide significant association mapped to *LINC01819* (also known as *KPRT4*; ACAT $P = 1.27\text{e-}6$), a long non-coding RNA that has been associated with some psychological phenotypes (e.g., insomnia⁴⁴) but has no known relevance to neurobiology.

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

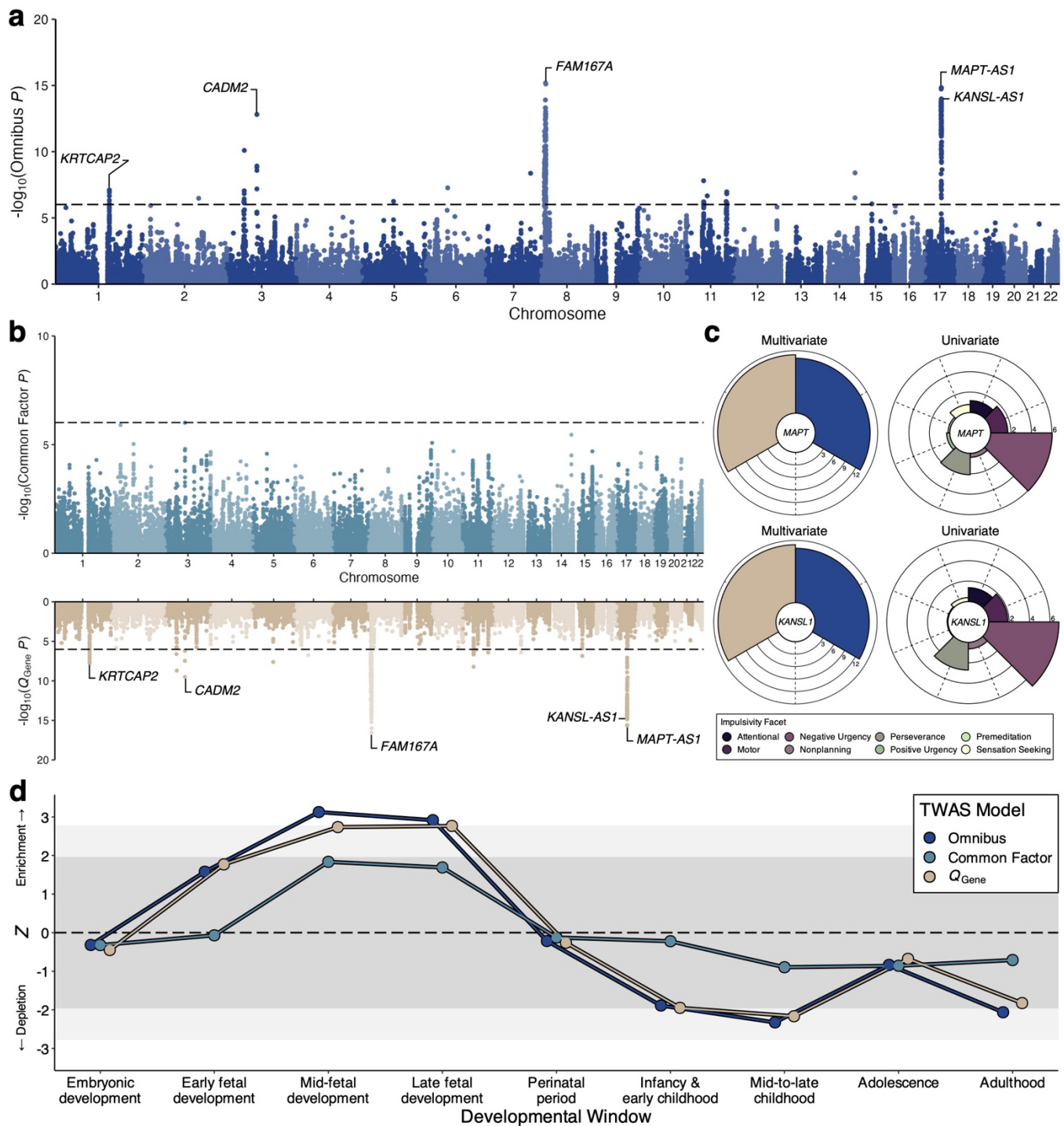


Figure 2. Multivariate modeling of transcriptomic architecture and its links to neurodevelopment. **a,b**, Manhattan (**a**) and Miami (**b**) plots of TWAS results for the omnibus and common factor models, respectively, where the x-axis refers to chromosomal position, the y-axis refers to statistical significance on a $-\log_{10}$ scale, the horizontal dashed line denotes genome-wide significance ($P = 9.67e-7$), and each point corresponds to one of the tested 51,698 gene-tissue pairs (spanning 13 GTEx brain tissues). **c**, Polar plots of ACAT TWAS P values for *MAPT* and *KANSL1* on a $-\log_{10}$ scale across all multivariate and univariate tests. **d**, Line plot of gene property analysis for TWAS results, illustrating patterns of enrichment and depletion of transcriptomic signals in developmentally specific genes across the lifespan. Developmental windows measured in PsychENCODE are plotted on the x-axis, while the Z statistic of enrichment or depletion is plotted on the y-axis. The dark gray band denotes the non-significant range of Z statistics, the light gray bands reflect nominal significance, and white reflects significance at a Bonferroni-corrected threshold.

Molecular bases of impulsivity facets implicate differing aspects of neurobiology

To characterize the biological pathways involved in the etiology of impulsivity, we used MAGMA to conduct gene-based association analyses for each of the three multivariate models (**Methods**). These gene-based test statistics were then used in gene property and gene set enrichment analyses to identify specific tissues, biological processes, molecular functions, and cellular components implicated in the molecular genetic architecture of impulsivity.

Consistent with the results of our multivariate GWAS, MAGMA-based association analyses implicated 64, 1, and 68 genes in the omnibus, common factor, and Q_{SNP} models, respectively (**Table S6**). As expected, all three models showed significant enrichment in brain tissues, with the omnibus model evincing the broadest and strongest pattern (**Figure S4, Table S7**). We also found that the omnibus GWAS results were enriched for genes preferentially expressed during fetal neurodevelopment ($P = 0.001$; **Figure S5, Table S8**). Follow-up facet-level analyses revealed similar results. For example, all impulsivity facets showed a broad pattern of enrichment across brain tissues, albeit to a weaker degree than the multivariate models (**Figure S4**). Similarly, developmental results indicated that genes preferentially expressed during early-to-mid fetal neurodevelopment tended to be most strongly enriched among facets. The exception to this pattern involved non-planning and urgency traits, which showed the greatest relative enrichment in genes preferentially expressed during late infancy (**Figure S5**).

With respect to gene set enrichment, no gene sets were significantly enriched in the omnibus or common factor results after correcting for multiple comparisons, though we did observe an enrichment of Q_{SNP} signal in genes related to synapse organization ($P = 3.39\text{e-}6$). Further inspection suggested there was a general pattern of enriched signal in gene sets related to synaptic biology across models, as many of the top-ranked gene sets were related to the synapse (**Table S9**). Subsequent permutation testing indicated that synaptic location and process gene sets were indeed more strongly enriched in all GWAS results compared to other collective sets of the same size (all $P_s < 1\text{e-}4$, **Figure S6; Methods**).

To characterize the biology implicated in the transcriptomic architecture of impulsivity, we translated the MAGMA statistical framework to facilitate gene-based association, gene property, and gene set enrichment analyses of multivariate TWAS results (**Methods**). Briefly, paralleling the results described above, we found that signal from the omnibus TWAS was enriched in genes specifically expressed during early to mid-fetal development at multiple stages ($P_s < 0.004$; **Figure 2d; Table S10**). In contrast, neither the common factor TWAS nor the Q_{Gene} results showed any pattern of enrichment across neurodevelopment. Facet-level results revealed that most impulsivity facets exhibited a similar pattern of enrichment as the omnibus model, except for premeditation, which showed a suggestive *depletion* of signal in genes specifically expressed during early prenatal development (**Figure S7**).

For TWAS-based gene set enrichment analyses, results implicated nitric oxide mediated signal transduction genes in the omnibus transcriptomic architecture of impulsivity ($P = 2.92\text{e-}8$; **Table S11**) – noteworthy given the key regulatory role of nitric oxide in central and peripheral neuronal signaling^{58,59}. Follow-

up tests revealed that this gene set was not significantly enriched in any of the lesser-powered facet-level results; however, there were nominally significant signals for attentional impulsivity and sensation seeking ($P_s = 0.003$ and 0.006 , respectively) that may be contributing to the omnibus signal. Finally, to parallel the MAGMA-based analyses, we performed permutation testing to evaluate whether gene sets related to synaptic locations and processes were enriched in the TWAS results, finding no evidence of categorical enrichment for these gene sets (all $P_s > 0.151$; **Methods**).

Impulsivity facets are differentially related to endophenotypes and other complex traits

Within the Genomic SEM framework, we next conducted a series of genetic correlation and Q_{Trait} analyses to explore how impulsivity relates to other complex traits and further evaluate the psychometric validity of the common impulsivity factor. Specifically, we used the common factor and Q_{Trait} models to estimate genetic relationships between impulsivity and three categories of traits: imaging-derived measures of brain structure, accelerometer-based metrics of physical activity, and psychological and psychiatric outcomes (**Methods**).

To gain greater insight into the etiological relationships between impulsivity and neuroanatomy, we estimated the genetic correlations between the common impulsivity factor and 34 bilateral regional measures of cortical thickness and surface area. The common factor was mostly negatively correlated with these measures of cortical morphology at a genetic level (**Figure 3a, Table S12**), which suggests that variants associated with a greater liability for impulsivity also tend to be associated with reduced cortical thickness ($r_g = -0.21$ to -0.019 [mean = -0.096]) and surface area ($r_g = -0.115$ to 0.063 [mean = -0.032]). Notably, some of the stronger genetic relationships between impulsivity and reduced cortical thickness were consistent with a shared pathway in frontotemporal regions (**Figure 3a,b,c**), including the lateral orbitofrontal cortex ($r_g = -0.210$, s.e. = 0.065 , $P = 0.001$, $Q_{\text{Trait}} P = 0.117$) and the medial orbitofrontal cortex ($r_g = -0.183$, s.e. = 0.069 , $P = 0.008$, $Q_{\text{Trait}} P = 0.084$).

Generally, though, Q_{Trait} analyses revealed widespread heterogeneity throughout the cortex for both thickness and surface area associations (**Figure 3c, Table S12**). With respect to cortical thickness, we observed strong evidence of trait-specific genetic relationships between impulsivity facets and cortical thickness, including in all three subregions of the inferior frontal gyrus: the pars orbitalis ($Q_{\text{Trait}} P = 1.3e-9$), pars triangularis ($Q_{\text{Trait}} P = 8.09e-11$), and pars opercularis ($Q_{\text{Trait}} P = 2.43e-7$). Notably, each of these regions has been functionally implicated in partially distinct aspects of self-regulatory control, as well as switching between goal-directed and habit-based responses. Examination of the facet-level genetic correlations revealed heterogeneity within the inferior frontal gyrus (**Figure 3d**), with Q_{Trait} signals driven by the divergent effects of attentional impulsivity and negative urgency, which, at times, had weak positive associations with cortical thickness. At the network level, topographical annotation indicated that cortical regions in the visual and default mode networks tended to harbor the most trait-specific surface area associations ($P_{\text{spin}} = 0.005$; **Figure 3e; Supplementary Information**).

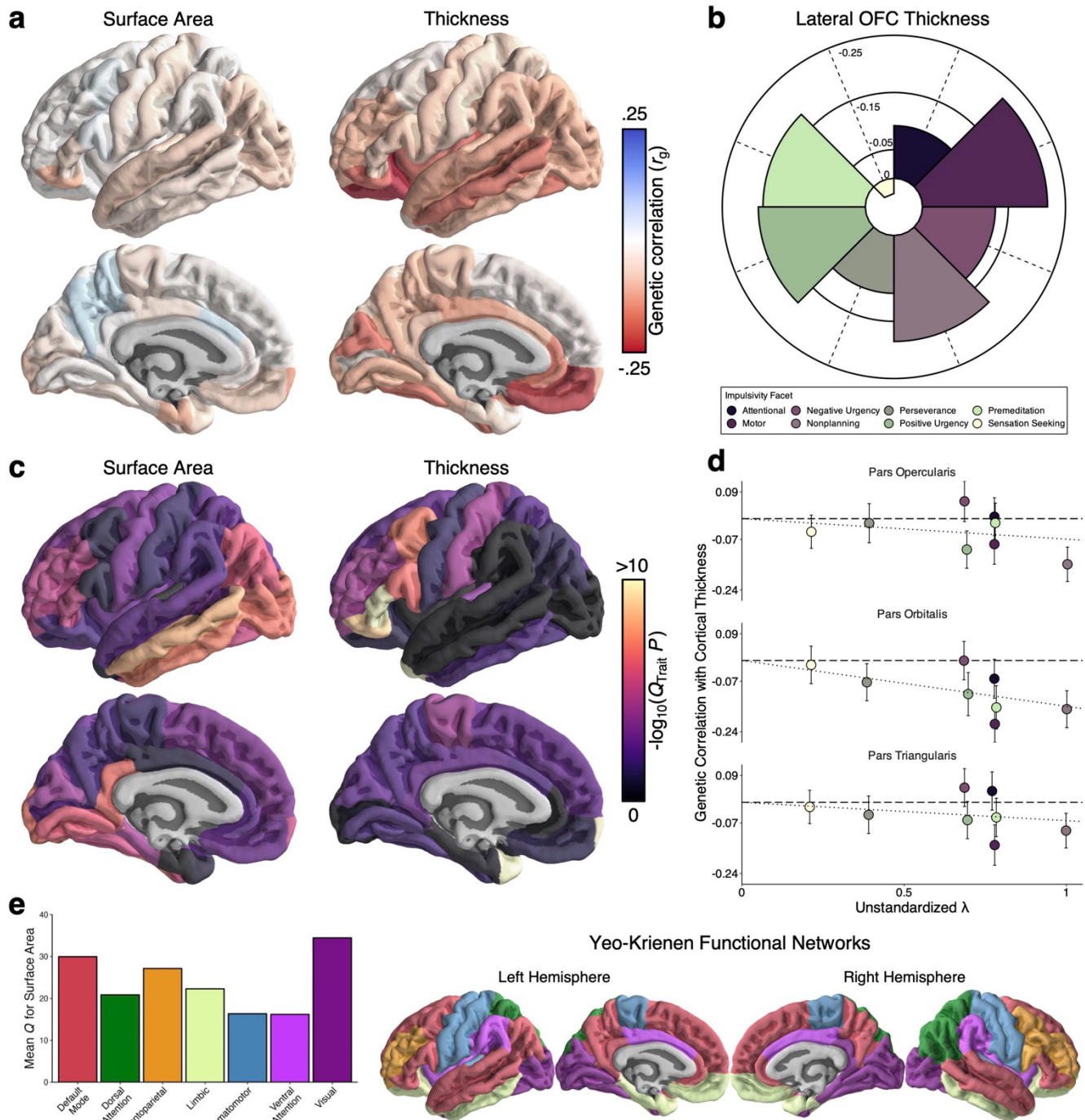


Figure 3. Topography of genetic relationships between cortical morphology and impulsivity. **a**, Map of genetic correlations between the common impulsivity factor and cortical surface area and thickness. **b**, Polar plot of genetic correlations between specific facets of impulsivity and thickness of the lateral orbitofrontal cortex, where facet-level effects are consistent with a shared pathway. **c**, Map of $Q_{\text{Trait}} -\log_{10}(P)$ values for cortical surface area and thickness, indexing the degree to which effects are incompatible with a common factor model. **d**, Scatter plots depicting the facet-level genetic relationships for all three subdivisions of the inferior frontal gyrus as a function of unstandardized factor loadings. **e**, Bar chart illustrating the difference in Q_{Trait} for cortical surface area across the Yeo-Krienen functional networks, which are displayed on the right for reference.

Given that motor alterations are a prominent feature of several neurological and psychiatric disorders, we next characterized the genomic relationships between impulsivity and accelerometer-based metrics of physical activity. We found that the common impulsivity factor was modestly genetically correlated with physical activity across the day, with positive genetic correlations during the early morning hours (00:00–06:59) and negative genetic correlations throughout the remainder of the day (07:00–23:59) (**Table S13; Figure S8**). Only 3 hourly associations were significant following correction for multiple testing, all of which showed significant evidence of heterogeneity in subsequent Q_{Trait} analyses. Inspection of the facet-level genetic correlations revealed that the links between impulsivity and physical activity were generally similar during morning hours, but these diverged in the afternoon and evening (**Figure 4a, Table S13**). Perseverance and sensation seeking displayed some of the clearest facet-specific patterns, with the former linked to a more pronounced dip in activity during the afternoon (lowest $r_g = -0.231$) and the latter linked to increased activity during the evening (highest $r_g = 0.178$).

Finally, we examined genomic relationships between impulsivity and 23 psychiatric phenotypes. As might be expected, genetic correlations between the common impulsivity factor and these outcomes were generally positive (range = 0.039 - 0.621) and statistically significant (**Figure 4b, Figure S9**). We observed some of the largest genetic correlations with attention-deficit/hyperactivity disorder ($r_g = 0.567$, s.e. = 0.043), alcohol use disorder ($r_g = 0.535$, s.e. = 0.072), and cannabis use disorder ($r_g = 0.479$, s.e. = 0.058). However, follow-up Q_{Trait} analyses revealed significant facet-level specificity for every tested trait (all $Q_{\text{Trait}} P < 8.48\text{e-}6$; **Figure 4c, Table S14; Figure S10**). For example, attention-deficit/hyperactivity disorder and its various clinical presentations (e.g., childhood, persistent, late-onset) exhibited null genetic correlations with perseverance in a manner that deviated from the common pathway model. Other notable deviations include facet-specific genetic correlations between schizophrenia and premeditation ($r_g = -0.104$, s.e. = 0.036), bipolar disorder and motor impulsivity ($r_g = 0.22$, s.e. = 0.039), and alcohol consumption and sensation seeking ($r_g = 0.228$, s.e. = 0.043) (**Figure 4c**). Collectively, results revealed that genetic relationships with psychopathology were inconsistent with a model of general liability to impulsivity and instead better captured with facet-specific pathways.

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

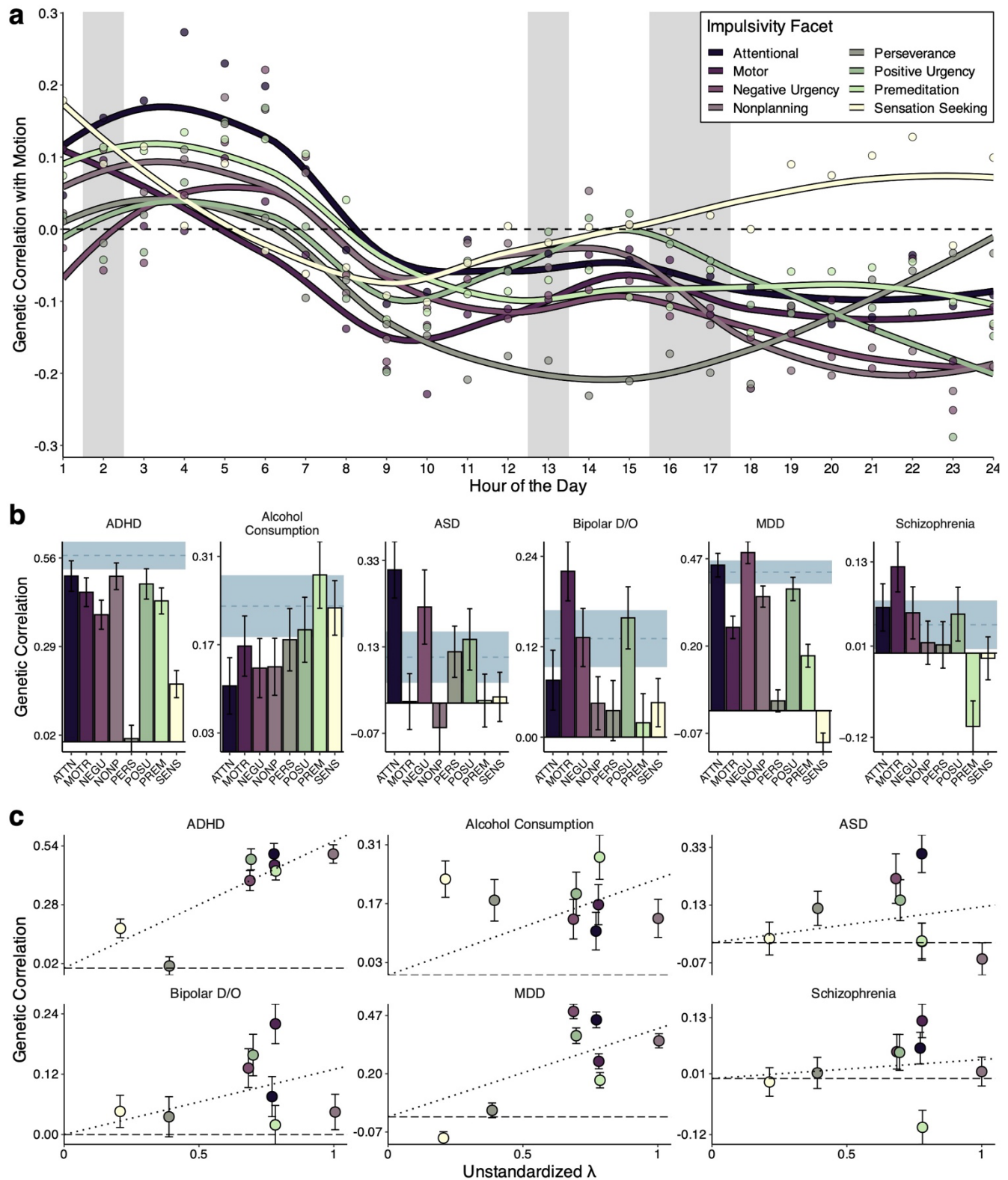


Figure 4. Facet-specific and heterogeneous genetic correlations between impulsivity and behavioral phenotypes. **a**, Line plots of genetic correlations between impulsivity facets and accelerometer-based measures of physical activity, as estimated with LOESS. Vertical gray bands highlight time periods when $Q_{\text{Trait}} P$ was non-significant (i.e., periods when the observed pattern of effects was consistent with a common pathway). **b**, Bar charts depicting the genetic correlations between impulsivity facets and select psychiatric outcomes. Horizontal blue lines index the genetic correlation between an outcome and the impulsivity factor (and the corresponding standard error). **c**, Scatter plots illustrating how genetic correlations with psychiatric outcomes are inconsistent with a common factor model in multiple ways, where facet-level relationships are plotted as a function of unstandardized factor loadings. Error bars depict standard errors.

Facet-level polygenic scores are generally more informative than one for a common factor of impulsivity

We next tested whether a polygenic score for common impulsivity might be more strongly or broadly associated with impulsivity and other health-related outcomes. To accomplish this, we calculated polygenic scores for the common factor of impulsivity and the constituent facets in two independent samples that span most of the human lifecourse: the Adolescent Brain and Cognitive Development Study (ABCD)^{60,61} and the Vanderbilt University Medical Center biobank (BioVU)^{62,63}. ABCD was used to examine associations between polygenic scores and psychological measures, while BioVU was used to examine associations with medical outcomes.

In ABCD, we found that the common factor polygenic score was robustly associated with psychological measures closely related to impulsivity, with similar ΔR^2 values as the facet-level polygenic scores across eight outcomes: the five UPPS-P subscales (negative urgency, [lack of] premeditation, [lack of] perseverance, positive urgency, and sensation seeking) and three subscales from the Child Behavior Checklist (CBCL)^{64,65} (attention, externalizing, and internalizing problems). Here, through a set of novel *Q*-like analyses, we found that the effects of facet-level polygenic scores largely aligned with a common pathway model (**Figure 5a, Methods**) – a finding that was further supported by the relatively strong performance of the impulsivity factor polygenic score in ABCD (**Table S15**). There were, however, six instances where the effects of sensation seeking or perseverance polygenic scores significantly deviated from expectation (**Table S15, Figure S11**), suggesting that some relationships in the sample are better captured by facet-specific versus shared pathways.

A very different pattern emerged in BioVU, where the common factor polygenic score was neither strongly nor broadly associated with health-related outcomes, with only five phenome-wide significant associations. By comparison, polygenic scores for negative urgency and nonplanning were linked to 52 and 47 medical outcomes, respectively. More precisely, when compared to all facet-level polygenic scores, significant associations with 67 unique health outcomes were obfuscated when the common factor polygenic score was used (**Figure 5b, Table S16, Figures S12-20**). Some of the missed relationships included novel links between urgency traits and increased odds for neurodevelopmental disorders (e.g., negative urgency–autism spectrum disorder: OR = 1.219, $P = 2e-8$; positive urgency–autism spectrum disorder: OR = 1.197, $P = 8.82e-5$; positive urgency–pervasive developmental disorder: OR = 1.119, $P = 1.64e-5$), negative urgency and increased odds for disorders of the circulatory system (e.g., negative urgency–myocardial infarction: OR = 1.103, $P = 1.57e-7$), and nonplanning and increased odds for viral hepatitis C (OR = 1.151, $P = 1.42e-6$). In addition to the expected associations with externalizing disorders (e.g., substance use disorders), there were significant positive links between multiple facets of impulsivity and internalizing conditions (e.g., major depressive disorder), buttressing the etiological connections between a broad spectrum of dysregulated behavioral and emotional processes.

Given the distinct PheWAS results, we sought to better understand the relationship between the common factor and facet-level polygenic score effects. In general, we found that the correlation between the two varied considerably across scores and medical categories (**Figure 5c**). Even in the mental disorders category, where concordance was generally higher, patterns ranged from modest negative correlations ($r = -0.213$ for perseverance) to strong positive correlations ($r = 0.771$ for nonplanning impulsivity).

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

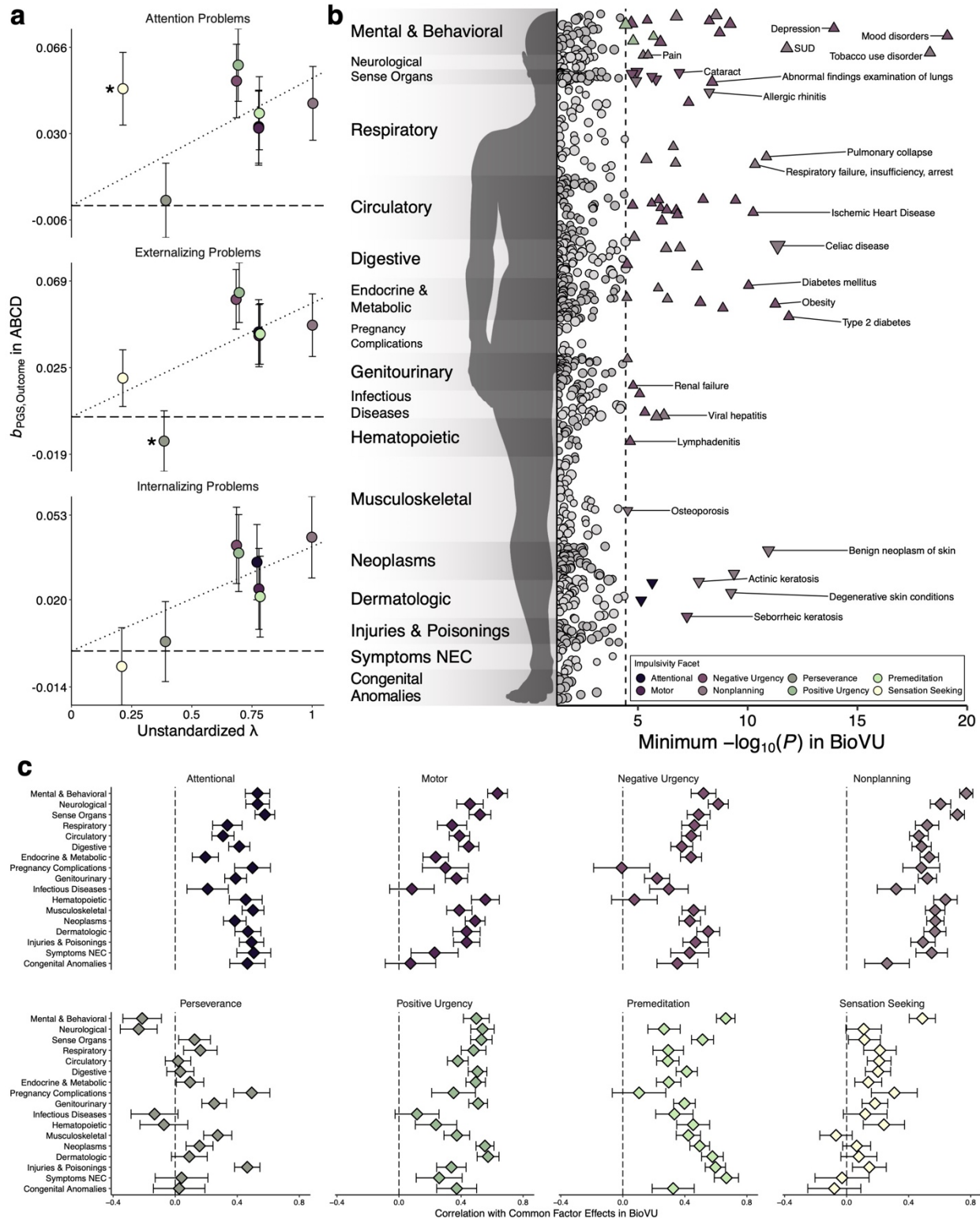


Figure 5. Polygenic scores and the validity of a common pathway in independent datasets. **a**, Scatter plot depicting the effect of each facet-level polygenic score for the CBCL outcomes in the ABCD Study as a function of the unstandardized factor loadings. Error bars depict the standard error. Asterisks denote significant deviations from the common pathways model. **b**, Manhattan plot of the strongest associations between facet-level polygenic scores and medical outcomes defined by “phecodes” in BioVU. The y-axis refers to the category of phecode, the x-axis refers to the statistical significance of the association on $-\log_{10}$ scale, and the dashed line denotes the Bonferroni-adjusted significance threshold ($P = 0.05/1,380 = 3.62e-5$). Each point in the plot corresponds to the minimum P value for an association between the eight facet-level polygenic scores and a phecode. Circles denote nonsignificant associations, upward-facing triangles are significant positive associations, and downward-facing triangles are significant negative associations. Size reflects the effect size, with larger points reflecting a larger absolute effect. For plotting purposes, the x-axis has been truncated to begin at $P < .05$ (1.301 on the $-\log_{10}$ scale), and conditions that did not meet this threshold were omitted. NEC = not elsewhere classified, SUD = substance use disorder. **c**, Scatter plot of the correlation between the effects of the common factor polygenic score and the effects of facet-level polygenic scores, where error bars correspond to the standard error of the estimate.

Discussion

Advancing etiological understanding of impulsivity is necessary to develop more effective therapeutic interventions for numerous psychiatric and neurological disorders. In the present study, we investigated the genetic architecture of impulsivity, employing complementary multivariate models to advance genomic discovery and formally evaluate pleiotropy across biological scales. Our results provide insight into the heterogeneous biology underlying impulsivity, revealing pervasive pleiotropy that is largely inconsistent with a common factor model of impulsivity in downstream analyses at molecular genetic, transcriptomic, neurogenomic, and phenomic levels. Three key insights emerge from this work: (1) there is more evidence for *splitting over lumping* that emerges at granular levels of analysis, especially for sensation seeking and (lack of) perseverance; (2) despite limited support for lumping, there are instances where the common factor approach may facilitate robust and interpretable discoveries for impulsivity; and (3) the etiology of impulsivity is closely linked to early neurodevelopment and related conditions.

The considerably greater evidence for fractionating impulsivity stands in stark contrast to recent findings in psychiatric and behavioral genetics, which indicate that substantial genetic risk is shared across related psychological outcomes. This is particularly striking since a common factor model showed good fit to our data, and this would typically suggest overlapping architecture that is consistent with a general dimension of liability. However, while a parsimonious lumping approach seemed viable at the level of genome-wide covariance, it performed poorly in subsequent analyses where many trait-specific patterns and signals were observed – especially for sensation seeking and (lack of) perseverance. These results align with related insights from neuroscience⁷, and they underscore the importance of evaluating conceptual models across levels of analysis, especially considering the growing trend to group related psychological constructs together^{13,14}.

Notably, our results show a distinct advantage of the omnibus model (i.e., the unstructured splitting approach) in GWAS and TWAS results. This approach markedly improved statistical power, enabling the discovery of novel risk loci and biological pathways related to multiple aspects of human impulsivity. For instance, our results revealed strong novel signals within chromosome 17 that localized to cytogenetic region 17q21.31. Here, we found a particularly robust association between impulsivity and the *MAPT* gene, along with its related non-coding RNA transcripts *MAPT-ASI* (antisense) and *MAPT-ITI* (intronic). Converging lines of evidence have linked genetic variation in *MAPT* to difficulties with self-regulation in the general population^{31,66}, with potential mechanisms including altered expression of either *MAPT* or its regulatory RNA transcripts, which can disrupt tau protein function⁶⁷. This disruption can lead to aberrant microtubule dynamics, synaptic dysfunction, and compromised neural circuitry involved in impulse control and executive functions^{68,69}. We also identified robust associations between impulsivity and *KANSL1* and its antisense transcript, *KANSL1-ASI*, which play a key role in autophagy and synaptic activity⁷⁰. Variation in this gene has previously been linked to neurodevelopmental and neurodegenerative disorders that exhibit difficulties with self-regulation^{53,56,71}, and changes to its expression or function could influence individual differences in impulsivity as well.

The improved statistical power of the omnibus model also facilitated new discoveries in downstream analyses, such as links to synaptic biology and nitric oxide signaling in the genomic and transcriptomic architecture of impulsivity, respectively. Nitric oxide plays a complex and multifaceted role in the central nervous system⁵⁹, where it is involved in mediating and modulating neurotransmission, neurovascular coupling, and neuroinflammatory processes. Disrupted nitric oxide metabolism also engenders impulsive, hyperactive, and aggressive behavior in murine models⁷²⁻⁷⁵, and altered nitric oxide signaling is linked to attention deficits in patients with certain neurometabolic disorders (e.g., argininosuccinate lyase deficiency)⁷⁶.

Our heterogeneity tests also advanced understanding of the pleiotropy across biological scales by identifying loci with divergent or trait-specific effects, such as *CADM2*. Here, we used Q_{SNP} and Q_{Gene} analyses to reveal that the effects of *CADM2* were highly specific to certain facets and inconsistent with a general liability to impulsivity, granting new insights not afforded by univariate GWAS or TWAS³⁸. Bioinformatic analyses indicated that the most heterogeneous and facet-specific signals localized to genes that are preferentially expressed during early neurodevelopment. Furthermore, our Q_{Trait} analyses detected several divergent effects on brain structure. We found that the inferior frontal gyrus exhibited pronounced heterogeneity with impulsivity, with highly facet-specific signals observed in all three of its cytoarchitecturally diverse subdivisions: the pars orbitalis, pars triangularis, and pars opercularis. As alterations in cortical thickness partly reflect variability in cellular organization (e.g., disturbances in neural progenitor cell proliferation or neuronal differentiation)^{77,78}, differences in cell type distributions or morphology may influence regional-specific links to impulse control and decision-making.

Beyond clarifying the unique biology associated with specific impulsivity facets, the Q tests also allowed us to identify effects that plausibly operate via a common pathway, such as the *FOXP2* gene. Q_{SNP} and Q_{Gene} results indicated that the effects of this gene on impulsivity were consistent with a general liability model, which is noteworthy given its prior associations with externalizing psychopathology^{31,48,49}, as well as speech and language disorders^{47,79,80}. With established roles in the regulation of neural plasticity, synaptic connectivity, and neurite outgrowth⁸¹, *FOXP2* may exert a highly pleiotropic effect on psychological outcomes during early development, with impulsivity serving as an intermediate phenotype. Furthermore, we found several nominal genetic correlations with cortical thickness that were consistent with a common pathway model, such as the medial and the lateral orbitofrontal cortex (OFC). Both of these regions have been functionally implicated in impulsive behavior, with medial OFC responses encoding the subjective value of delayed reward in order to support goal-directed behavior⁸², and lateral OFC responses, together with the insula, supporting learning from aversive outcomes to guide adaptive behavior⁸³⁻⁸⁵.

Finally, we identified several specific risk genes and pathways that are involved in critical developmental programs, and we found that both the genetic and transcriptomic architectures of impulsivity are enriched for genes that are primarily expressed during early-to-mid fetal neurodevelopment. These links were underscored in our polygenic score analyses, including the first PheWAS of impulsivity, where we identified novel associations between facet-level polygenic scores and neurodevelopmental conditions. In scanning the medical phenome, we

found that genetic liability for urgency-related traits was linked to autism spectrum disorder in a manner that appeared to be trait-specific, with effects that were larger than any other association (e.g., negative urgency–autism spectrum disorder: OR = 1.219; positive urgency–autism spectrum disorder: OR = 1.197). Notably, these relationships were not observed when using a common factor polygenic score. However, although the common factor polygenic score was not informative in adults, we did find that it was robustly associated with psychological measures in children, where facet-level effects largely aligned with a common pathway model. One interpretation of this pattern is that genetic risk for impulsivity largely operates via broad, general pathways in children but canalizes into more unique channels across development, such that a general index becomes less informative by adulthood, as seen for problematic alcohol use^{86–88}.

These findings should be interpreted in light of several limitations, such as the exclusion of individuals with non-European ancestry due to methodological constraints. This, coupled with the participation biases in the present cohort⁸⁹, may affect the generalizability of our findings to other populations. Furthermore, all univariate GWAS analyses were conducted using facet-level sum scores from the BIS-11 and the UPPS-P, as opposed to item-level measures. It has previously been demonstrated that the use of item-level modeling can ameliorate certain confounds and generate deeper insights into the shared genetic architecture among complex traits^{30,66,90}. Additionally, the present study was limited to self-report measures of impulsivity, which may not fully generalize to task-based measures of similar cognitive and behavioral processes. Thus, future research should seek to replicate and extend our findings in more diverse and representative samples with more granular measures of impulsive traits.

Nevertheless, our study considerably advances understanding of impulsive traits and their genetic architectures. Through multivariate modeling across biological scales, we identified novel risk genes for impulsivity facets, uncovered trait-specific effects at multiple levels of analysis, and highlighted new genetic links with neurodevelopment. While we do find several instances of convergent biology that implicate a shared pathway, we find substantially more evidence for trait-specific biology underlying the impulsivity facets. Sensation seeking and perseverance clearly differ from other facets in their etiological underpinnings, and they are perhaps better conceptualized as indices of psychological motivation or reward sensitivity rather than self-regulation, *per se*³⁷. Indeed, this would align with phenotypic factor analytic research that suggests these facets tend to be more distinct from other impulsive traits^{91,92}.

Collectively, these results cast considerable doubt on the validity of impulsivity as a unitary construct, challenging the conventional practice of lumping together diverse cognitive and behavioral tendencies that share only modest overlap. Instead, our findings advocate for a more refined and nuanced approach to studying impulsivity and its links to psychiatric and neurological disorders, one that recognizes its multifaceted nature and the distinct biology that influences its various facets. This shift in perspective paves the way for a deeper exploration of impulsive traits, promising new insights into the complexities of human psychology and health.

Methods

This study is accompanied by **Supplementary Information**, which includes additional tables and figures.

GWAS data acquisition and pre-processing

Univariate GWAS summary statistics were obtained from a previously published study of impulsivity³⁸. In the present study, we analyzed GWAS summary statistics for two psychometrically validated measures of impulsivity: the Barratt Impulsiveness Scale (BIS-11)³⁹ and the UPPS-P Impulsive Behavior Scale⁴⁰. The former includes subscales that measure attentional, motor, and nonplanning impulsiveness, while the latter includes subscales that measure (lack of) premeditation, (lack of) perseverance, positive urgency, negative urgency, and sensation seeking. These phenotypes are succinctly described below:

- **Attentional Impulsiveness (ATTN)**: An inability to concentrate or focus attention.
- **Motor Impulsiveness (MOTR)**: A predisposition to hastily engage in behaviors in response to internal or external stimuli.
- **Negative Urgency (NEGU)**: A tendency to react rashly to negative affect without consideration of potential consequences.
- **Nonplanning Impulsiveness (NONP)**: An inclination to forgo future planning and forethought.
- **Lack of Perseverance (PERS)**: An inability to remain focused on tasks that may be boring or difficult.
- **Positive Urgency (POSU)**: A tendency to rashly react to positive affect without consideration of potential consequences.
- **Lack of Premeditation (PREM)**: A tendency not to think about or reflect on the consequences of an action before taking it.
- **Sensation Seeking (SENS)**: A predisposition to seek out new and exciting experiences, potentially without evaluating the risk.

As described by the original authors³⁸, these GWAS were conducted in a cohort of up to 133,517 participants of European ancestry collected by 23andMe, Inc., a direct-to-consumer genetics company. Further description of the cohort can be found in the original study by Sanchez-Roige and colleagues³⁸. All summary statistics included in the present analysis had significant SNP-based heritability and robust polygenic signal (**Table 1**). Reference alleles were aligned across univariate GWAS summary statistics, and only SNPs with a minor allele frequency $\geq 0.5\%$ and an imputation quality score ≥ 0.6 were included in analyses.

Multivariate GWAS of impulsivity

Using Genomic SEM v0.0.5^{29,33}, we performed multivariate genome-wide association analyses for two distinct statistical models: the omnibus model (i.e., “splitting” impulsivity) and the common factor model (i.e.,

“lumping” impulsivity). We also perform follow-up Q_{SNP} tests to determine whether SNP effects on the common factor are statistically plausible. Details of the exploratory and confirmatory factor analyses are reported in the **Supplementary Information**. All analyses were conducted adhering to best practices, using standard reference panels and parameter settings, and diagonally weighted least squares (DWLS) estimation. The estimated sample size of the latent genetic factor was calculated in the manner described by Mallard and colleagues³².

We refer interested readers to the original methodology papers by Grotzinger and colleagues^{29,33} for exhaustive description of the method, but we provide an overview of the genomic structural equation models reported herein below. As illustrated in **Figure 1c,d**, the omnibus model tests the degree to which a SNP has non-zero effects on impulsivity facets, the common factor model tests the degree to which a SNP influences impulsivity facets via a common pathway, and the Q_{SNP} model tests the degree to which said common pathway is statistically plausible.

Omnibus model. To test whether the overall pattern of effects for a given SNP was non-zero across study phenotypes, we performed genome-wide unstructured tests of association, where all relationships among impulsivity facets and their relationships with tested SNPs were freely estimated. This was accomplished by first fitting the following system of regressions for each SNP:

$$\begin{aligned}y_1 &= b_1 \cdot \text{SNP}_i + \epsilon_1 \\y_2 &= b_2 \cdot \text{SNP}_i + \epsilon_2 \\&\vdots \\y_k &= b_k \cdot \text{SNP}_i + \epsilon_k\end{aligned}$$

Where y refers to the k facets in the structural equation model, SNP_i is the tested SNP, b represents the estimated regression coefficients, and ϵ denotes the error terms. We then fit a null model with all regression coefficients constrained to equal zero for the tested SNP, such that

$$\begin{aligned}y_1 &= 0 \cdot \text{SNP}_i + \epsilon_1 \\y_2 &= 0 \cdot \text{SNP}_i + \epsilon_2 \\&\vdots \\y_k &= 0 \cdot \text{SNP}_i + \epsilon_k\end{aligned}$$

The difference in fit between these two models was then calculated with a χ^2 difference test, where $\Delta\chi^2 = \chi^2_{\text{free}} - \chi^2_{\text{constrained}}$ and $\Delta df = df_{\text{free}} - df_{\text{constrained}}$. The omnibus P value for SNP_i can then be calculated as $P = 1 - CDF_{\chi^2}(\Delta\chi^2, \Delta df)$, where CDF_{χ^2} is the cumulative distribution function of the chi-square distribution with Δdf degrees of freedom.

Common factor model. To estimate the effect of a given SNP on a general dimension of impulsivity, we used the common factor approach described by Grotzinger and colleagues²⁹ to model the genetic covariances

among study phenotypes. In this framework, a measurement model describes the genetic relationships between model indicators and a latent factor as

$$\begin{aligned} y_1 &= \lambda_1 \cdot F + \epsilon_1 \\ y_2 &= \lambda_2 \cdot F + \epsilon_2 \\ &\vdots \\ y_k &= \lambda_k \cdot F + \epsilon_k \end{aligned}$$

Where y refers to the k observed indicators in the measurement model, F is the latent genetic factor, λ represents the factor loadings, and ϵ denotes the residual variances or errors associated with each indicator. The latent genetic factor can then be regressed onto a given SNP as

$$F = b \cdot SNP_i + \zeta$$

Here, b refers to the regression coefficient representing the effect of SNP_i on F , and ζ is the residual. Note that the standard error and P value for this estimated parameter were computed within the model, as well.

Q_{SNP} model. To statistically test whether the estimated SNP effect plausibly operated via the latent factor (i.e., a common pathway), we conducted follow-up Q_{SNP} tests for every tested SNP. This involved comparing the common factor model to one where SNPs directly influence indicators through freely estimated independent pathways, as opposed to the common pathway. The difference in model fit can then be quantified with a χ^2 difference test, where $\Delta\chi^2 = \chi^2_{\text{independent}} - \chi^2_{\text{common}}$ and $\Delta df = df_{\text{independent}} - df_{\text{common}}$. The Q_{SNP} P value for SNP_i was subsequently calculated as $P = 1 - CDF_{\chi^2}(\Delta\chi^2, \Delta df)$, as described above.

Note that the expectation is not necessarily that SNP-level effects are homogenous across model indicators. Instead, under the common pathway model, we would expect that the effect of a given SNP on model indicators scales proportionally with the unstandardized factor loadings, where

$$\mathbb{E} \begin{bmatrix} b_{SNP_i, y_1} \\ b_{SNP_i, y_2} \\ \vdots \\ b_{SNP_i, y_k} \end{bmatrix} = b_{SNP_i, F} \cdot \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{bmatrix}$$

Here, b_{SNP_i, y_k} represents the effect of SNP i on indicator k , $b_{SNP_i, F}$ is the effect of the SNP i on latent factor F , and λ_k represents the unstandardized loadings on the latent factor. As described by de la Fuente and colleagues⁹³, misfit will occur when the vector of expected SNP effects on indicators is discordant with the observed SNP effect on indicators. The resulting Q statistic will be high, and the model in which the SNP

influences indicators exclusively via the common pathway will be rejected. Interpretation can be facilitated by plotting the univariate GWAS betas by the unstandardized factor loadings, as illustrated in **Figure 1g,h**.

Biological annotation of GWAS results

To characterize genome-wide significant loci, we used the FUMA v1.5.4 SNP2GENE pipeline with default parameters to apply a standard clumping algorithm to our GWAS results. The European ancestry samples from the 1000 Genomes Project Phase 3v5 ($n = 503$) were used as a LD reference panel. To identify putative risk genes, we conducted positional mapping using field standard ANNOVAR annotations (version 2017.07.17). We also used FUMA to investigate whether loci associated with impulsivity have been previously linked to other complex traits in the GWAS Catalog (version e0_r2022-11-29).

Gene-based analyses. We used MAGMA⁹⁴ v1.08 to conduct gene-based association, gene property, and gene set enrichment analyses. Standard procedures were followed for gene-based association analyses based on GWAS summary statistics and default MAGMA parameters were employed throughout the pipeline.

MAGMA was also used to conduct competitive gene property and gene-set enrichment analyses based on the gene-level P values produced in the association analyses. These analyses evaluated whether genes in a specific annotation had a stronger association with the GWAS outcome compared to other genes.

In the gene property analyses, we utilized the GTEx v8 dataset, comprising 54 tissues, to examine enrichment in various bodily tissues. Additionally, we used brain tissue data from 11 developmental epochs in the BrainSpan dataset to assess enrichment throughout neurodevelopment. In the gene set enrichment analyses, we tested up to 9,988 gene sets from the Molecular Signatures Database v2023.1. These sets correspond to 7,343 biological processes, 1,001 cellular components, and 1,644 molecular functions. For determining statistical significance, we used Bonferroni-corrected thresholds of $P \leq 9.26e-4$ for the GTEx tests, $P \leq 4.54e-3$ for the BrainSpan tests, and $P \leq 5.01e-6$ for the gene set tests.

Permutation test for synaptic gene sets. Given the relatively strong signals observed for several gene sets related to the synapse, we sought to test whether gene sets robustly implicated in synaptic biology, especially those robustly linked to synaptic locations and processes, were significantly enriched in the genetic architecture of impulsivity. To accomplish this, we utilized the SynGO database⁹⁵, a meticulously curated repository of synaptic gene ontologies, to identify with high confidence gene sets related to the synapse as well as those that were not. We then used a permutation-based approach to test whether the mean Z statistic from our synapse-related gene sets was significantly different from what we might expect by chance, which we empirically established by randomly sampling the mean Z statistic of 10,000 combinations of gene sets of the same size. Statistical significance was calculated based on the number of times the observed value was greater than the randomly sampled value (i.e., $P = \frac{\text{number of permuted means} \geq \text{observed mean}}{\text{total number of permutations}}$).

Multivariate TWAS of impulsivity

To investigate the effects of tissue-specific gene expression on impulsivity, we used Transcriptome-Wide Structural Equation Modeling⁹⁶ (version 0.0.5), which is based on the FUSION TWAS method. Specifically, we used this approach to conduct multivariate transcriptome-wide association analyses for the omnibus and common factor models. We also performed follow-up tests and calculated Q_{Gene} statistics to evaluate whether SNP effects on the common factor were plausible.

Within this framework, GWAS summary statistics for the impulsivity facets were then paired with *cis* gene expression quantitative trait locus data for 13 brain tissues from the Genotype-Tissue Expression Project (GTEx) v8 dataset to produce univariate TWAS estimates for 51,698 gene-tissue pairs. These gene-level summary statistics were then used as input for the multivariate TWAS, which yielded results for the omnibus, common factor, and Q_{Gene} models.

Aggregated Cauchy association test. Given that effects were strongly and positively correlated across tissues (all cross-tissue $r_s > 0.856$), we used an aggregated Cauchy association test (ACAT) to integrate information across all available tissues and improve statistical power prior to downstream bioinformatic analyses. Specifically, for each multivariate model, we used the ACAT R package (version 0.91) to combine TWAS P values across available tissues for each gene. Following the recommendations of the developers, any ACAT P value equal to 1 was replaced by $1-1/d$, where d refers to the number of TWAS P values combined by ACAT. We calculated Z statistics using the inverse cumulative density function (or quantile function) of the standard normal distribution, such that $Z = \Phi^{-1}(1 - \frac{P}{2})$.

Biological annotation of TWAS results

To assess the enrichment of TWAS signals across various biological contexts, we employed a series of regression-based procedures, drawing inspiration from the MAGMA framework⁹⁴. Specifically, we examined enrichment across neurodevelopmental epochs and gene ontology terms.

For neurodevelopmental enrichment, we leveraged gene expression data from the PsychENCODE Consortium, deriving gene expression specificity scores for nine neurodevelopmental epochs (see following subsection). Each epoch's enrichment was then assessed by regressing gene-level Z statistics, obtained from the ACAT TWAS P values, on specificity scores for each developmental window. The median number of SNPs used in the TWAS weights, the median gene expression heritability, and the number of TWAS P values combined via ACAT were included as covariates in this analysis.

For gene ontology terms, we characterized the enrichment of TWAS signals for all terms in the Molecular Signatures Database v2023.1. This was achieved by regressing the ACAT TWAS Z statistics onto a binary variable indicating gene set membership while accounting for the same gene-level covariates described above. Gene sets with less than three genes available for analysis were excluded. Finally, to parallel our GWAS bioinformatic analyses, we also conducted permutation-based tests of enrichment for SynGO⁹⁵ gene sets as described in the previous section.

Statistical significance was tested using two-sided P values for the continuous neurodevelopmental specificity scores, while one-sided P values were used for the binary gene sets, consistent with the MAGMA framework. Bonferroni-corrected thresholds for significance were set at $P \leq 5.55e-3$ for developmental analyses and $P \leq 7.72e-6$ for gene set analyses.

Specificity scores. To assess the specificity of expression for each gene across development, we calculated specificity scores in PsychENCODE data, using a similar specificity metric as Skene and colleagues⁹⁷. First, we employed a normalization procedure that scaled the expression of a gene in each developmental window based on its relative expression in other windows, using the developmental periods defined by PsychENCODE investigators⁹⁸. This was achieved by computing the average expression of each gene within each window as $A_{wj} = \frac{1}{n_w} \sum_{i=1}^{n_w} E_{ij}$, where A_{wj} denotes the average expression of gene j for window w , E_{ij} represents the expression level of gene j in sample i from window w , and n_w is the number of samples associated with window w . We then calculated the specificity score, S , for each gene in each window as $S_{wj} = \frac{A_{wj}}{\sum_{k=1}^m A_{kj}}$, where m is the total number of windows.

Genetic correlations with other complex traits

Genomic SEM v0.0.5, which is based on LD score regression^{99,100}, was used to estimate genetic correlations between general impulsivity and other complex traits, as well as between the impulsivity facets and these traits. In the present paper, we selected traits from three domains for study: (i) imaging-derived measures of cortical structure, (ii) accelerometer-based measures of physical activity throughout the day, and (iii) psychiatric outcomes.

Genetic correlations were estimated adhering to best practices, using standard reference panels and parameter settings. All GWAS summary statistics were processed with the *munge* function of Genomic SEM, which retained only HapMap3 SNPs outside of the major histocompatibility complex regions with an allele frequency ≥ 0.01 for analysis. Within each family of tests, P values were adjusted for multiple comparisons via a Bonferroni correction. Note that all genetic correlations with the common factor were subsequently examined with Q_{Trait} tests, as previously described³³. This test parallels Q_{SNP} and Q_{Gene} , as it evaluates the degree to which genetic relationships between model indicators and an external trait plausibly operate via a common pathway, aligning with the magnitude of factor loadings.

Polygenic score analyses

PRS-CS¹⁰¹ v2021.06.04 and PLINK¹⁰² v2.00 were used to calculate polygenic scores for the impulsivity facets and the common factor of impulsivity. Briefly, we first used PRS-CS to apply a continuous shrinkage prior to SNP effect estimates and infer posterior SNP weights for 875,517 common SNPs (minor allele frequency ≥ 0.01) that were present across the HapMap3 and 1000 Genomes Project Phase 3v5 datasets. European ancestry samples from the 1000 Genomes Project Phase 3v5 reference panel were used to model LD among variants. We

then used PLINK to calculate polygenic scores for individuals in each sample by summing all included variants weighted by the inferred posterior effect size for the effect allele. The resulting polygenic scores were subsequently standardized to unit variance.

This process was used to calculate polygenic scores in two independent datasets: the Adolescent Brain and Cognitive Development Study (ABCD)^{60,61} and the Vanderbilt University Medical Center Biobank (BioVU)^{62,63}. The former was used to examine associations between genetic liability for impulsivity and a select number of psychological outcomes in children, while the latter was used to examine associations between genetic liability for impulsivity and a wide array of medical outcomes, as described below.

Adolescent Brain and Cognitive Development Study. We used the ABCD Study ($n = 4,314$) to examine relationships between liability for impulsivity and closely related psychological outcomes in children. Specifically, we sought to characterize associations between impulsivity polygenic scores and eight outcomes measured at baseline: the five UPPS-P subscales (negative urgency, [lack of] premeditation, [lack of] perseverance, positive urgency, and sensation seeking)⁴⁰ and three subscales from the Child Behavior Checklist (CBCL) (attention, externalizing, and internalizing problems)^{64,65}.

To accomplish this, we used the lavaan R package (version 0.6-12) to conduct multiple regression analyses with DWLS estimation. For each outcome, we first fit a baseline model that only included sex, age, and the first 10 principal components of ancestry as covariates. We then fit models that iteratively added each polygenic score to the model, freely estimating the effect of the polygenic score on the outcome while adjusting for covariates. The variance explained by each score was calculated as $\Delta R^2 = R^2_{\text{full}} - R^2_{\text{covariates}}$. As results revealed that the common factor polygenic score performed quite well in terms of ΔR^2 , we next sought to determine if the pattern of effects was consistent with a common pathway model. Below, we describe a novel means of testing for facet-specific relationships that is conceptually similar to Q_{SNP} , Q_{Gene} , and Q_{Trait} tests, but allows us to test whether individual effects deviate from expectation.

Recall that facet-level effects are expected to scale proportionally with the magnitude of unstandardized factor loadings if the common pathway model is valid. Thus, we began by establishing the expected beta coefficient for each facet-level polygenic score based on the effect of the common factor polygenic score and the facet's unstandardized factor loading from the genomic factor model, such that

$$\mathbb{E} \begin{bmatrix} b_{PGS_1, outcome} \\ b_{PGS_2, outcome} \\ \vdots \\ b_{PGS_k, outcome} \end{bmatrix} = b_{PGS_{CF}, outcome} \cdot \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{bmatrix}$$

We then fit a follow-up model with a constraint imposed, wherein the effect of the polygenic score was fixed to its expected value. The freely estimated and constrained models were then compared using a likelihood ratio test, which allowed us to quantitatively determine whether deviations from expectation were statistically significant. Note that this approach is only suitable when the univariate GWAS (i.e., the input for Genomic SEM),

have comparable statistical power. That is to say, if polygenic scores for the model indicators vary drastically in their predictive power, we may observe deviations that are not informative about the common pathway.

Vanderbilt University Medical Center Biobank. The genotyped BioVU sample ($n = 66,915$) served as the basis for examining associations between impulsivity polygenic scores and a broad range of medical phenotypes. The genotyping process and the quality control measures for this sample have been described in previous publications. Medical phenotypes, termed here as “phecodes”, were derived from the International Classification of Disease (ICD) diagnostic codes present in the electronic health records of participants. For a participant to be recognized as a case for a given phecode, we required that two separate ICD diagnostic code instances be present in their record. Only phecodes with at least 100 cases were considered for analysis. This resulted in 1,380 phecodes being included in the present study.

The PheWAS R package (version 0.99.5-2) was then used to conduct phenome-wide association analyses, where, for each of the phecodes, a logistic regression model was used to estimate the odds of each diagnosis based on each impulsivity polygenic score. Covariates included sex, median age from the longitudinal electronic health record data, and the first 10 principal components of ancestry. Phenome-wide significance was set at a Bonferroni-adjusted threshold of $P \leq 3.62e-5$.

Acknowledgments

We would like to thank the research participants and employees of 23andMe for making this work possible. The following members of the 23andMe Research Team contributed to this study: Stella Aslibekyan, Adam Auton, Elizabeth Babalola, Robert K. Bell, Jessica Bielenberg, Jonathan Bowes, Katarzyna Bryc, Ninad S. Chaudhary, Daniella Coker, Sayantan Das, Emily DelloRusso, Sarah L. Elson, Nicholas Eriksson, Teresa Filshtein, Pierre Fontanillas, Will Freyman, Zach Fuller, Chris German, Julie M. Granka, Karl Heilbron, Alejandro Hernandez, Barry Hicks, David A. Hinds, Ethan M. Jewett, Yunxuan Jiang, Katelyn Kukar, Alan Kwong, Yanyu Liang, Keng-Han Lin, Bianca A. Llamas, Matthew H. McIntyre, Steven J. Micheletti, Meghan E. Moreno, Priyanka Nandakumar, Dominique T. Nguyen, Jared O'Connell, Aaron A. Petrakovitz, G. David Poznik, Alexandra Reynoso, Shubham Saini, Morgan Schumacher, Leah Selcer, Anjali J. Shastri, Janie F. Shelton, Jingchunzi Shi, Suyash Shringarpure, Qiaojuan Jane Su, Susana A. Tat, Vinh Tran, Joyce Y. Tung, Xin Wang, Wei Wang, Catherine H. Weldon, Peter Wilton, Corinna D. Wong.

We would also like to add that the current analyses would not have been possible without the enormous efforts put forth by the investigators and participants from the Adolescent Brain Cognitive Development Study, BioVU, ENIGMA, iPSYCH, PsychENCODE, Psychiatric Genomics Consortium, and UK Biobank.

Author Contributions

T.T.M. and S.S-R. conceived the study. The study protocol was developed by T.T.M. with input from S.S-R. T.T.M. was the lead analyst responsible for this study with assistance from J.D.T., M.J., and Y.Z. T.T.M. drafted the manuscript, with substantive contributions from J.W.S., A.A.P., and S.S-R. T.T.M. prepared the figures and tables with assistance from J.D.T., M.J., Y.Z., M.L.W., and C.M.W. Various aspects of study design and framing benefitted from advice and feedback from D.E.G., A.D.G., M.L.W., R.G.F., L.K.D., A.R., E.M.T-D., K.W.C., and T.G. All authors critically reviewed the manuscript and contributed to its final draft.

Competing Interests

P.F. and S.E. are employed by and hold stock or stock options in 23andMe, Inc. J.W.S. is a member of the Leon Levy Foundation Neuroscience Advisory Board, the Scientific Advisory Board of Sensorium Therapeutics (with equity), and PI of a collaborative study of the genetics of depression and bipolar disorder sponsored by 23andMe, Inc. for which the company provides analysis time as in-kind support but no payments.

Data Availability

No new data were collected for this study. Qualified researchers may download the impulsivity GWAS summary statistics through 23andMe, Inc. via a data use agreement that protects the privacy of their participants. Please visit <https://research.23andme.com/collaborate/#dataset-access/> for more information regarding data access.

References

1. Evenden, J. L. Varieties of impulsivity. *Psychopharmacology* **146**, 348–361 (1999).
2. Moeller, F. G., Barratt, E. S., Dougherty, D. M., Schmitz, J. M. & Swann, A. C. Psychiatric Aspects of Impulsivity. *AJP* **158**, 1783–1793 (2001).
3. Robbins, T. W., Gillan, C. M., Smith, D. G., Wit, S. de & Ersche, K. D. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends in Cognitive Sciences* **16**, 81–91 (2012).
4. Dalley, J. W., Everitt, B. J. & Robbins, T. W. Impulsivity, Compulsivity, and Top-Down Cognitive Control. *Neuron* **69**, 680–694 (2011).
5. Kim, S. & Lee, D. Prefrontal Cortex and Impulsive Decision Making. *Biological Psychiatry* **69**, 1140–1146 (2011).
6. Rosenbaum, G. M. & Hartley, C. A. Developmental perspectives on risky and impulsive choice. *Philosophical Transactions of the Royal Society B: Biological Sciences* **374**, 20180133 (2018).
7. Dalley, J. W. & Robbins, T. W. Fractionating impulsivity: neuropsychiatric implications. *Nat Rev Neurosci* **18**, 158–171 (2017).
8. Vrieze, S. I., McGue, M., Miller, M. B., Hicks, B. M. & Iacono, W. G. Three Mutually Informative Ways to Understand the Genetic Relationships Among Behavioral Disinhibition, Alcohol Use, Drug Use, Nicotine Use/Dependence, and Their Co-occurrence: Twin Biometry, GCTA, and Genome-Wide Scoring. *Behav Genet* **43**, 97–107 (2013).
9. Mann, F. D. *et al.* Sensation seeking and impulsive traits as personality endophenotypes for antisocial behavior: Evidence from two independent samples. *Personality and Individual Differences* **105**, 30–39 (2017).
10. Fortgang, R. G., Hultman, C. M., Erp, T. G. M. van & Cannon, T. D. Multidimensional assessment of impulsivity in schizophrenia, bipolar disorder, and major depressive disorder: testing for shared endophenotypes. *Psychological Medicine* **46**, 1497–1507 (2016).
11. Krueger, R. F. *et al.* Etiologic connections among substance dependence, antisocial behavior and personality: Modeling the externalizing spectrum. *J Abnorm Psychol* **111**, 411–424 (2002).
12. Sanchez-Roige, S. *et al.* Genome-Wide Association Studies of Impulsive Personality Traits (BIS-11 and UPPS-P) and Drug Experimentation in up to 22,861 Adult Research Participants Identify Loci in the CACNA1I and CADM2 genes. *J Neurosci* **39**, 2562–2572 (2019).
13. Insel, T. *et al.* Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *American Journal of Psychiatry* (2010).
14. Kotov, R. *et al.* The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology* **126**, 454 (2017).
15. McGorry, P. D., Hartmann, J. A., Spooner, R. & Nelson, B. Beyond the “at risk mental state” concept: transitioning to transdiagnostic psychiatry. *World Psychiatry* **17**, 133–142 (2018).
16. Friedman, N. P. & Gustavson, D. E. Do Rating and Task Measures of Control Abilities Assess the Same Thing? *Curr Dir Psychol Sci* **31**, 262–271 (2022).
17. Gray, J. A. The psychophysiological basis of introversion-extraversion. *Behaviour Research and Therapy* **8**, 249–266 (1970).
18. Gray, J. A. *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. 548 (Clarendon Press/Oxford University Press, 1982).
19. Bijttebier, P., Beck, I., Claes, L. & Vandereycken, W. Gray’s Reinforcement Sensitivity Theory as a framework for research on personality–psychopathology associations. *Clinical Psychology Review* **29**, 421–430 (2009).

20. Steinberg, L. A social neuroscience perspective on adolescent risk-taking. *Developmental Review* **28**, 78–106 (2008).
21. Steinberg, L. A dual systems model of adolescent risk-taking. *Developmental Psychobiology* **52**, 216–224 (2010).
22. Harden, K. P. & Tucker-Drob, E. M. Individual differences in the development of sensation seeking and impulsivity during adolescence: Further evidence for a dual systems model. *Developmental Psychology* **47**, 739–746 (2011).
23. Yip, S. W. & Potenza, M. N. Application of Research Domain Criteria to childhood and adolescent impulsive and addictive disorders: Implications for treatment. *Clinical Psychology Review* **64**, 41–56 (2018).
24. Ringwald, W. R., Forbes, M. K. & Wright, A. G. C. Meta-analysis of structural evidence for the Hierarchical Taxonomy of Psychopathology (HiTOP) model. *Psychological Medicine* **53**, 533–546 (2023).
25. Kotov, R. *et al.* Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): I. Psychosis superspectrum. *World Psychiatry* **19**, 151–172 (2020).
26. Krueger, R. F. *et al.* Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): II. Externalizing superspectrum. *World Psychiatry* **20**, 171–193 (2021).
27. Watson, D. *et al.* Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): III. Emotional dysfunction superspectrum. *World Psychiatry* **21**, 26–54 (2022).
28. Lee, P. H. *et al.* Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell* **179**, 1469–1482. e11 (2019).
29. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour* **3**, 513–525 (2019).
30. Thorp, J. G. *et al.* Symptom-level modelling unravels the shared genetic architecture of anxiety and depression. *Nat Hum Behav* **5**, 1432–1442 (2021).
31. Karlsson Linnér, R. *et al.* Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. *Nat Neurosci* **24**, 1367–1376 (2021).
32. Mallard, T. T. *et al.* Multivariate GWAS of psychiatric disorders and their cardinal symptoms reveal two dimensions of cross-cutting genetic liabilities. *Cell Genom* **2**, (2022).
33. Grotzinger, A. D. *et al.* Genetic architecture of 11 major psychiatric disorders at biobehavioral, functional genomic and molecular genetic levels of analysis. *Nat Genet* **54**, 548–559 (2022).
34. Grotzinger, A. D. Shared genetic architecture across psychiatric disorders. *Psychological Medicine* **51**, 2210–2216 (2021).
35. Mallard, T. T., Grotzinger, A. D. & Smoller, J. W. Examining the shared etiology of psychopathology with genome-wide association studies. *Physiol Rev* (2023) doi:10.1152/physrev.00016.2022.
36. Waszczuk, M. A. *et al.* Dimensional and transdiagnostic phenotypes in psychiatric genome-wide association studies. *Mol Psychiatry* 1–11 (2023) doi:10.1038/s41380-023-02142-8.
37. Gustavson, D. E. *et al.* Executive function and impulsivity predict distinct genetic variance in internalizing problems, externalizing problems, thought disorders, and compulsive disorders: a genomic structural equation modeling study. *Clinical Psychological Science* (in press).
38. Sanchez-Roige, S. *et al.* CADM2 is implicated in impulsive personality and numerous other traits by genome- and phenome-wide association studies in humans and mice. *Transl Psychiatry* **13**, 1–11 (2023).
39. Patton, J. H., Stanford, M. S. & Barratt, E. S. Factor structure of the barratt impulsiveness scale. *Journal of Clinical Psychology* **51**, 768–774 (1995).
40. Cyders, M. A. *et al.* Integration of impulsivity and positive mood to predict risky behavior: Development and validation of a measure of positive urgency. *Psychological Assessment* **19**, 107–118 (2007).
41. Howard, D. M. *et al.* Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience* **22**, 343–352 (2019).

42. Okbay, A. *et al.* Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nat Genet* **54**, 437–449 (2022).
43. Saunders, G. R. B. *et al.* Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature* **612**, 720–724 (2022).
44. Watanabe, K. *et al.* Genome-wide meta-analysis of insomnia prioritizes genes associated with metabolic and psychiatric pathways. *Nat Genet* **54**, 1125–1132 (2022).
45. Nagel, M. *et al.* Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nature Genetics* **50**, 920–927 (2018).
46. Baselmans, B. M. L. *et al.* Multivariate genome-wide analyses of the well-being spectrum. *Nat Genet* **51**, 445–451 (2019).
47. Lai, C. S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **413**, 519–523 (2001).
48. Tielbeek, J. J. *et al.* Uncovering the genetic architecture of broad antisocial behavior through a genome-wide association study meta-analysis. *Mol Psychiatry* **27**, 4453–4463 (2022).
49. Demontis, D. *et al.* Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains. *Nat Genet* **55**, 198–208 (2023).
50. Steinberg, K. M. *et al.* Structural diversity and African origin of the 17q21.31 inversion polymorphism. *Nat Genet* **44**, 872–880 (2012).
51. Boettger, L. M., Handsaker, R. E., Zody, M. C. & McCarroll, S. A. Structural haplotypes and recent evolution of the human 17q21.31 region. *Nat Genet* **44**, 881–885 (2012).
52. Koolen, D. A. *et al.* A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. *Nat Genet* **38**, 999–1001 (2006).
53. Koolen, D. A. *et al.* Mutations in the chromatin modifier gene KANSL1 cause the 17q21.31 microdeletion syndrome. *Nat Genet* **44**, 639–641 (2012).
54. Zollino, M. *et al.* Mutations in KANSL1 cause the 17q21.31 microdeletion syndrome phenotype. *Nat Genet* **44**, 636–638 (2012).
55. Simón-Sánchez, J. *et al.* Genome-wide association study reveals genetic risk underlying Parkinson’s disease. *Nat Genet* **41**, 1308–1312 (2009).
56. Nalls, M. A. *et al.* Identification of novel risk loci, causal insights, and heritable risk for Parkinson’s disease: a meta-analysis of genome-wide association studies. *The Lancet Neurology* **18**, 1091–1102 (2019).
57. Wang, H. *et al.* Chromosomal inversion polymorphisms shape human brain morphology. *Cell Reports* **42**, (2023).
58. Garthwaite, J. NO as a multimodal transmitter in the brain: discovery and current status. *British Journal of Pharmacology* **176**, 197–211 (2019).
59. Lundberg, J. O. & Weitzberg, E. Nitric oxide signaling in health and disease. *Cell* **185**, 2853–2878 (2022).
60. Casey, B. J. *et al.* The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev Cogn Neurosci* **32**, 43–54 (2018).
61. Uban, K. A. *et al.* Biospecimens and the ABCD study: Rationale, methods of collection, measurement and early data. *Developmental Cognitive Neuroscience* **32**, 97–106 (2018).
62. Roden, D. M. *et al.* Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clinical Pharmacology & Therapeutics* **84**, 362–369 (2008).
63. Dennis, J. K. *et al.* Clinical laboratory test-wide association scan of polygenic scores identifies biomarkers of complex disease. *Genome Medicine* **13**, 6 (2021).
64. Achenbach, T. M. The classification of children’s psychiatric symptoms: A factor-analytic study. *Psychological Monographs: General and Applied* **80**, 1–37 (1966).
65. Achenbach, T. M. & Ruffle, T. M. The Child Behavior Checklist and Related Forms for Assessing Behavioral/Emotional Problems and Competencies. *Pediatrics In Review* **21**, 265–271 (2000).

66. Mallard, T. T. *et al.* Item-Level Genome-Wide Association Study of the Alcohol Use Disorders Identification Test in Three Population-Based Cohorts. *Am J Psychiatry* **179**, 58–70 (2022).
67. Caillet-Boudin, M.-L., Buée, L., Sergeant, N. & Lefebvre, B. Regulation of human MAPT gene expression. *Molecular Neurodegeneration* **10**, 28 (2015).
68. Wang, Y. & Mandelkow, E. Tau in physiology and pathology. *Nat Rev Neurosci* **17**, 22–35 (2016).
69. Zhang, Y., Wu, K.-M., Yang, L., Dong, Q. & Yu, J.-T. Tauopathies: new perspectives and challenges. *Molecular Neurodegeneration* **17**, 28 (2022).
70. Linda, K. *et al.* Imbalanced autophagy causes synaptic deficits in a human model for neurodevelopmental disorders. *Autophagy* **18**, 423–442 (2022).
71. Koolen, D. A. *et al.* The Koolen-de Vries syndrome: a phenotypic comparison of patients with a 17q21.31 microdeletion versus a KANSL1 sequence variant. *Eur J Hum Genet* **24**, 652–659 (2016).
72. Chiavegatto, S. *et al.* Brain serotonin dysfunction accounts for aggression in male mice lacking neuronal nitric oxide synthase. *Proceedings of the National Academy of Sciences* **98**, 1277–1281 (2001).
73. Nelson, R. J., Trainor, B. C., Chiavegatto, S. & Demas, G. E. Pleiotropic contributions of nitric oxide to aggressive behavior. *Neuroscience & Biobehavioral Reviews* **30**, 346–355 (2006).
74. Tanda, K. *et al.* Abnormal social behavior, hyperactivity, impaired remote spatial memory, and increased D1-mediated dopaminergic signaling in neuronal nitric oxide synthase knockout mice. *Molecular Brain* **2**, 19 (2009).
75. Gao, Y. & Heldt, S. A. Lack of neuronal nitric oxide synthase results in attention deficit hyperactivity disorder-like behaviors in mice. *Behavioral Neuroscience* **129**, 50–61 (2015).
76. Lerner, S. *et al.* ASL Metabolically Regulates Tyrosine Hydroxylase in the Nucleus Locus Coeruleus. *Cell Reports* **29**, 2144–2153.e7 (2019).
77. Rakic, P. Specification of Cerebral Cortical Areas. *Science* **241**, 170–176 (1988).
78. Rakic, P. Evolution of the neocortex: a perspective from developmental biology. *Nat Rev Neurosci* **10**, 724–735 (2009).
79. MacDermot, K. D. *et al.* Identification of FOXP2 Truncation as a Novel Cause of Developmental Speech and Language Deficits. *The American Journal of Human Genetics* **76**, 1074–1080 (2005).
80. Vargha-Khadem, F., Gadian, D. G., Copp, A. & Mishkin, M. FOXP2 and the neuroanatomy of speech and language. *Nat Rev Neurosci* **6**, 131–138 (2005).
81. Fisher, S. E. & Scharff, C. FOXP2 as a molecular window into speech and language. *Trends in Genetics* **25**, 166–177 (2009).
82. Kable, J. W. & Glimcher, P. W. The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* **10**, 1625–1633 (2007).
83. O’Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J. & Andrews, C. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* **4**, 95–102 (2001).
84. Wächter, T., Lungu, O. V., Liu, T., Willingham, D. T. & Ashe, J. Differential Effect of Reward and Punishment on Procedural Learning. *J. Neurosci.* **29**, 436–443 (2009).
85. Elliott, R., Agnew, Z. & Deakin, J. F. W. Hedonic and Informational Functions of the Human Orbitofrontal Cortex. *Cerebral Cortex* **20**, 198–204 (2010).
86. Kendler, K. S., Gardner, C. & Dick, D. M. Predicting alcohol consumption in adolescence from alcohol-specific and general externalizing genetic risk factors, key environmental exposures and their interaction. *Psychological Medicine* **41**, 1507–1516 (2011).
87. Meyers, J. L. *et al.* Genetic Influences on Alcohol Use Behaviors Have Diverging Developmental Trajectories: A Prospective Study Among Male and Female Twins. *Alcoholism: Clinical and Experimental Research* **38**, 2869–2877 (2014).
88. Barr, P. B. *et al.* Parsing genetically influenced risk pathways: genetic loci impact problematic alcohol use via externalizing and specific risk. *Transl Psychiatry* **12**, 1–8 (2022).

89. Pirastu, N. *et al.* Genetic analyses identify widespread sex-differential participation bias. *Nat Genet* **53**, 663–671 (2021).
90. Nagel, M., Watanabe, K., Stringer, S., Posthuma, D. & van der Sluis, S. Item-level analyses reveal genetic heterogeneity in neuroticism. *Nat Commun* **9**, 905 (2018).
91. Whiteside, S. P., Lynam, D. R., Miller, J. D. & Reynolds, S. K. Validation of the UPPS impulsive behaviour scale: a four-factor model of impulsivity. *Eur J Pers* **19**, 559–574 (2005).
92. Watts, A. L., Smith, G. T., Barch, D. M. & Sher, K. J. Factor structure, measurement and structural invariance, and external validity of an abbreviated youth version of the UPPS-P Impulsive Behavior Scale. *Psychological Assessment* **32**, 336–347 (2020).
93. de la Fuente, J., Davies, G., Grotzinger, A. D., Tucker-Drob, E. M. & Deary, I. J. A general dimension of genetic sharing across diverse cognitive traits inferred from molecular data. *Nature Human Behaviour* **5**, 49–58 (2021).
94. Leeuw, C. A. de, Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Computational Biology* **11**, e1004219 (2015).
95. Koopmans, F. *et al.* SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. *Neuron* **103**, 217-234.e4 (2019).
96. Grotzinger, A. D., de la Fuente, J., Davies, G., Nivard, M. G. & Tucker-Drob, E. M. Transcriptome-wide and stratified genomic structural equation modeling identify neurobiological pathways shared across diverse cognitive traits. *Nat Commun* **13**, 6280 (2022).
97. Skene, N. G. *et al.* Genetic identification of brain cell types underlying schizophrenia. *Nat Genet* **50**, 825–833 (2018).
98. Li, M. *et al.* Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, (2018).
99. Bulik-Sullivan, B. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291–295 (2015).
100. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236 (2015).
101. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nature Communications* **10**, 1–10 (2019).
102. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, s13742-015-0047–8 (2015).