The Genetic and Neural Substrates of Externalizing Behavior

Bart Baselmans, Anke R. Hammerschlag, Stephany Noordijk, Hill Ip, Matthijs van der Zee, Eco de Geus, Abdel Abdellaoui, Jorien L. Treur, and Dennis van ’t Ent

ABSTRACT

BACKGROUND: To gain more insight into the biological factors that mediate vulnerability to display externalizing behaviors, we leveraged genome-wide association study summary statistics on 13 externalizing phenotypes.

METHODS: After data classification based on genetic resemblance, we performed multivariate genome-wide association meta-analyses and conducted extensive bioinformatic analyses, including genetic correlation assessment with other traits, Mendelian randomization, and gene set and gene expression analyses.

RESULTS: The genetic data could be categorized into disruptive behavior (DB) and risk-taking behavior (RTB) factors, and subsequent genome-wide association meta-analyses provided association statistics for DB and RTB (N_{eff} = 523,150 and 1,506,537, respectively), yielding 50 and 257 independent genetic signals. The statistics of DB, much more than RTB, signaled genetic predisposition to adverse cognitive, mental health, and personality outcomes. We found evidence for bidirectional causal influences between DB and substance use behaviors. Gene set analyses implicated contributions of neuronal cell development (DB/RTB) and synapse formation and transcription (RTB) mechanisms. Gene-brain mapping confirmed involvement of the amygdala and hypothalamus and highlighted other candidate regions (cerebellar dentate, cuneiform nucleus, claustrum, paracentral cortex). At the cell-type level, we noted enrichment of glutamatergic neurons for DB and RTB.

CONCLUSIONS: This bottom-up, data-driven study provides new insights into the genetic signals of externalizing behaviors and indicates that commonalities in genetic architecture contribute to the frequent co-occurrence of different DBs and different RTBs, respectively. Bioinformatic analyses supported the DB versus RTB categorization and indicated relevant biological mechanisms. Generally similar gene-brain mappings indicate that neuroanatomical differences, if any, escaped the resolution of our methods.

https://doi.org/10.1016/j.bpsgos.2021.09.007
Identiﬁed factors using N-weighted genome-wide association studies (GWAS) of 13 externalizing phenotypes, we applied linkage disequilibrium score regression (LDSC) (24,25) to compute pairwise genetic correlations. Initially, we collected GWAS summary statistics for 13 externalizing phenotypes (Table 1). Only samples with a European/North-American ancestry (23) were included (see Supplement 1).

### Phenotype Selection

Based on the disinhibited externalizing spectrum described by Krueger et al. (1), we collected GWAS summary statistics for 13 externalizing phenotypes (Table 1). Only samples with a European/North-American ancestry (23) were included (see Supplement 1).

### Identiﬁcation of Genetic Factor Structures

To identify genetic factor structures among the 13 included phenotypes, we ﬁrst applied linkage disequilibrium score regression (LDSC) (24,25) to compute pairwise genetic correlations (r_p) (see Supplement 1). Second, knowing the genetic correlations, we examined relationships using hierarchical clustering with (1 − r_p) as genetic distance measures between phenotypes and with linkage based on Ward’s method (26). We followed the Calinski-Harabasz criterion to indicate the optimal number of clusters (27).

Third, in addition to hierarchical clustering, we applied factor analysis on the GWAS summary statistics using genomic structural equation modeling (28) (see Supplement 1).

### Multivariate N-Weighted Genome-wide Association Meta-analysis

Univariate GWAS summary statistics were separately combined for identiﬁed factors using N-weighted genome-wide association meta-analysis (GWAMA) (22), which is robust against sample overlap (see Supplement 1).

### Genetic Relationships With Other Traits

We computed pairwise genetic correlations using LDSC (24,25) for our identiﬁed factors with 61 additional phenotypes in the following categories: mental health, cognition and socioeconomic status, personality, social, substance use, cardiovascular disease risk, physical health, anthropomorphic, and reproduction (see Table S6 in Supplement 2 and Supplement 1 for details).

### Mendelian Randomization

For one of the identiﬁed clusters characterized by DBs, we applied MR to test for causal relationship with substance use behaviors (see Supplement 1).

### Tissue-Type Associations

We used the gene associations of the two factors as input for a tissue-type analysis using MAGMA version 1.08 (http://ctg.cncr.nl/software/magma) (29) with the N-weighted GWAMA summary statistics of the identiﬁed factors as input. The gene test statistics are deﬁned as the mean single nucleotide polymorphism (SNP) association using the sum of −log (SNP p value).

### Stratified LDSC of Tissue Types

We applied stratified LDSC to investigate which tissues and cell types are enriched for the identiﬁed factors (see Supplement 1).
RESULTS

Identification of Genetic Factor Structure

LDSC (Figure 1A) indicated substantial genetic correlations, particularly among phenotypes characterized by disruptive-type behaviors (e.g., aggression and ADHD: \( r_g = 0.72 \), SE = 0.18) and phenotypes characterized by risk-taking behaviors (RTBs) (e.g., cannabis use and number of sexual partners: \( r_g = 0.69 \), SE = 0.02), with less genetic overlap between pairs of phenotypes characterized by the different behavior types (e.g., aggression and cannabis use: \( r_g = 0.03 \), SE = 0.13). Hierarchical clustering on genetic resemblance (Figure 1B) confirmed this categorization into DBs (aggression, angry outbursts, different measures of irritability, and ADHD) and RTBs (antisocial behavior, general risk tolerance, drinks per week, ever smoker, number of sexual partners, automobile speeding propensity, and lifetime cannabis use). Exploratory factor analysis by genomic structural equation modeling supported a division into two phenotypic clusters by demonstrating a substantial increase in explained variance from 34% for a one-factor model to 53% for a two-factor model, but a much more substantial increase in explained variance from 34% for a one-factor model to 58% for a three-factor model (Figure 1C; Table S1 in Supplement 2). For RTB, we identified 257 independent genome-wide significant SNPs at 194 genomic risk loci (\( N_{	ext{eff}} = 1,506,537 \)) (Figure 2A; Table S4 in Supplement 2). The LD score intercepts were close to 1 for both DB (intercept = 1.0167, SE = 0.0086; LDSC ratio = 0.0367, SE = 0.0156) and RTB (intercept = 1.0031, SE = 0.0132; LDSC ratio = 0.0242, SE = 0.0106), indicating that neither population stratification nor sample overlap, but rather an increase of polygenic signal, was driving the SNP associations. The SNP heritability as defined by LDSC was 0.0396 (SE = 0.0015) for DB and 0.022 (SE = 0.0007) for RTB. The genetic correlation between DB and RTB was 0.33 (SE = 0.02).

Multivariate GWAMA

For both factors, we subsequently meta-analyzed the phenotype GWAS data using N-weighted GWAMA. For the DB factor, we identified 50 independent genome-wide significant SNPs at 42 loci (\( N_{	ext{eff}} = 523,150 \)) (Figure 2A; Table S4 in Supplement 2). For RTB, we identified 257 independent genome-wide significant SNPs at 194 genomic risk loci (\( N_{	ext{eff}} = 1,506,537 \)) (Figure 2B; Table S5 in Supplement 2). For these phenotypes, we followed the hierarchical clustering results, assigning ADHD to factor 1 (DB) and antisocial behavior to factor 2 (RTB).

Genetic Relationships With Other Traits

Figure 3 and Table S7 in Supplement 2 show pairwise genetic correlations for DB and RTB with 61 additional phenotypes. DB, much more than RTB, showed a pattern of genetic overlap pointing to adverse outcomes of cognition, socioeconomic status, several mental and physical health measures, and

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Genetic factor structure of externalizing behaviors. (A) Genetic correlations between the externalizing phenotypes calculated by linkage disequilibrium score regression. (B) Hierarchical clustering dendrogram based on genetic resemblance of the externalizing phenotypes. Blue represents the disruptive behavior cluster, and red represents the risk-taking behavior cluster. (C) Multivariate GWAMA results of the exploratory factor analysis of the genetic correlations presented in panel (A), a confirmatory factor model with two correlated genetic factors was specified using genomic structural equation modeling. In this model, the common factors account for the genetic covariation among the externalizing traits, i.e., each of the two common genetic factors represents variation in genetic liability that is shared across the phenotypes that load on it. Disruptive behavior represents shared genetic liability among disorders characterized by disruptive behavior, and risk-taking behavior represents the shared liability for risk-taking behavior. One-headed arrows connecting the common genetic factors to the individual traits represent standard loadings, which can be interpreted as coefficients from a regression of the true genetic liability for the trait on the common factor. Two-headed arrows connecting the genetic components represent their correlations. Two-headed arrows connecting the genetic components of the individual traits to themselves represent residual genetic variances and correspond to the proportion of heritable variation in liability to each individual trait that is unexplained by the two factors. ADHD, attention-deicit/hyperactivity disorder; Agg, aggression; Ao, angry outbursts; As, antisocial behavior; Asp, automobile speeding propensity; Can, lifetime cannabis use; Dr, drinks per week; Ir, extreme irritable; Ir2d, irritable for 2 days; R, general risk tolerance; Sm, ever smoker; Sx, number of sexual partners.
personality. For example, DB showed a negative genetic correlation with educational attainment ($r_g = -0.34, SE = 0.02$) and income ($r_g = -0.44, SE = 0.04$), whereas genetic correlations for RTB with these phenotypes were close to zero. In addition, DB was more positively correlated with depressive symptoms ($r_g = 0.73, SE = 0.03$) and neuroticism ($r_g = 0.68, SE = 0.04$) and more negatively correlated with agreeableness ($r_g = -0.59, SE = 0.05$) compared with RTB (depressive symptoms: $r_g = 0.21, SE = 0.03$; neuroticism: $r_g = -0.09, SE = 0.005$; agreeableness: $r_g = -0.12, SE = 0.05$).

Causal Relationship of DB With Smoking and Alcohol Use

We tested for bidirectional causal effects between DB (which does not include substance use phenotypes) and measures of smoking (33) and alcohol use (34) using MR. We focused on these specific relationships because DBs and substance use are particularly strongly associated, and knowledge of (the direction of) potential causal effects has major public health implications. When DB was the exposure variable, inverse variance weighted analyses provided strong evidence for causal effects such that DB increases the odds of smoking initiation ($\beta = 0.39, 95\% CI = 0.24–0.54, p = 2.0 \times 10^{-5}$) and decreases the odds of being able to successfully quit smoking ($\beta = -0.20, 95\% CI = -0.33 \text{ to } -0.07, p = .002$) (Table 2; Figure S1 in Supplement 1). Effect sizes and statistical evidence for a causal relationship between DB and smoking initiation were broadly consistent across the different MR methods. However, for DB to smoking cessation, the weighted mode and MR-Egger did not support a causal relationship (Table 2). Some evidence for pleiotropic effects was provided by a significant Cochran’s heterogeneity test (Cochran’s Q $p$ value = $2.4 \times 10^{-33}$) (Table S8 in Supplement 2) but not supported by the MR-Egger intercept (only DB $\rightarrow$ smoking initiation) (Table S9 in Supplement 2). Steiger filtering and MR-PRESSO did not affect the results and continued to support causal effects (Tables S10–S12 in Supplement 2). No consistent results were found for a causal relationship between DB and number of cigarettes smoked per day (Table 2; Table S9 in Supplement 2).

From DB to alcohol use disorder, there was evidence for a causal increasing effect (inverse variance weighted analysis: $\beta = 0.26, 95\% CI = 0.08–0.44, p = .004$), which was consistent (but weaker) in direction of effect across multiple MR methods (Table 2). MR-Egger was not available owing to low reliability; however, there was no heterogeneity based on Cochran’s Q ($p = .457$) (Table S8 in Supplement 2). Steiger filtering and MR-PRESSO did not change the outcome (Tables S10–S12 in Supplement 2).

In the opposite direction, we found strong evidence for a causal, increasing effect of smoking initiation on DB (inverse variance weighted analysis: $\beta = 0.17, 95\% CI = 0.14–0.20, p = 9.9 \times 10^{-31}$), which was consistent across weighted median, weighted mode, and generalized summary data–based MR (MR-Egger not available). Steiger filtered analyses and MR-
Figure 3. Genetic correlations with other phenotypes for disruptive behavior (blue) and risk-taking behavior (green) factors. Points represent the correlation estimates, and lines represent the 95% confidence intervals. Significant associations, after correction for multiple testing, are marked by orange stars. ADHD, attention-deficit/hyperactivity disorder; BMI, body mass index; HDL, high-density lipoprotein; LDL, low-density lipoprotein; MDD, major depressive disorder; SES, socioeconomic status.

PRESSO did not change the results (Table 2; Tables S10–S12 in Supplement 2). However, there was marked evidence for heterogeneity (Cochran’s Q p = 7.3 × 10^{-46}) (Table S8 in Supplement 2). There was no clear evidence for causal effects of alcohol use on DB.

Biological Annotations

Gene Associations. Genome-wide gene-based association analysis identified 81 genes significantly associated with DB (Table S13 in Supplement 2) and 318 genes significantly associated with RTB (Table S14 in Supplement 2) after Bonferroni correction. CADM2 showed an exceptionally strong association (2.10 × 10^{-29}) with RTB.

Tissue-Type Associations. We used the gene-based test statistics as input for gene set analysis in MAGMA. Starting with a tissue-type analysis, we found significant associations at the Bonferroni level for nine brain regions for DB and 13 brain regions for RTB (Tables S15 and S16 in Supplement 2). Conditional analyses indicated that only the frontal cortex showed an independent association with DB, while independent associations were limited to the frontal cortex and cerebellum for RTB.

Gene Set Associations. Subsequent gene set analyses identified three gene sets for DB involved in cell development and 13 gene sets for RTB involved in cell development, synapse formation, and transcription (Tables S17 and S18 in Supplement 2). Conditional gene set analyses on the three gene sets associated with DB showed that the three gene set associations were highly related (Table S17 in Supplement 2). In addition, the conditional analyses of the 13 gene sets identified for RTB traits revealed redundancy for multiple gene sets related to synapse and cell development and transcription, leaving seven gene sets with independent association (Table S18 in Supplement 2).

Stratified LDSC. Complementary to gene set analysis, we applied stratified LDSC to study effects on 10 general tissue-type groups. Consistent with the gene set analyses, these analyses indicated significant enrichment after Bonferroni correction of the central nervous system (CNS) for both DB (Z_{CNS} = 6.84) (Table S19 in Supplement 2) and RTB (Z_{CNS} = 7.97) (Table S20 in Supplement 2). It has to be noted that other tissues were associated with DB and RTB as well (e.g., liver, kidney), but these associations were considerably weaker (Tables S19 and S20 in Supplement 2). Zooming in on associations on the cell-type level for DB revealed 45 significant cell type–specific annotations after Bonferroni correction, of which the top 15 all involved the CNS (Table S21 in Supplement 2). A similar pattern was observed for RTB, where we found 96 significant cell type–specific annotations, of which 32 indicated CNS involvement (Table S22 in Supplement 2).
Table 2. Results of the Two-Sample Bidirectional Mendelian Randomization Analyses Between DB and Smoking and Alcohol Use Behavior

| Exposure | Outcome                  | SNPs, n | F    | β    | 95% CI     | p      | β     | 95% CI     | p      | β     | 95% CI     | p      | β     | 95% CI     | p     | β     | 95% CI     | p     |
|----------|--------------------------|---------|------|------|------------|--------|-------|------------|--------|-------|------------|--------|-------|------------|--------|-------|------------|--------|
|          |                          | IVW     | Weighted Median | Weighted Mode | MR-Egger (SIMEX) | GSMR   |
|          |                          |         |                  |              |            |        |                  |        |                  |        |                  |        |
| DB       | Smoking initiation       | 40      | 37.41            | 0.39       | 0.24       | 2.0 × 10⁻⁷ | 0.25   | 0.13       | 3.2 × 10⁻⁵ | 0.15   | -0.06       | 1.80 | 0.43 | 0.09       | .17    | 33    | 0.342      | 0.27 | 1.7 × 10⁻²⁰ |
|          |                          |         |                  |              |            |        |                  |        |                  |        |                  |        |
|          | Cigarettes/day           | 38      | 37.61            | 0.14       | 0.01       | 9.9    | 0.01   | 0.02       | 8.0    | 0.15   | 0.05       | .044 | 0.18 | 0.02       | .294   | 34    | 0.178      | 0.09  | 3.3 × 10⁻³⁵ |
|          |                          |         |                  |              |            |        |                  |        |                  |        |                  |        |
|          | Smoking cessation        | 40      | 37.33            | -0.20      | -0.33      | -0.07  | -0.11 | -0.25       | -0.28  | 0.02   | -0.14       | .17   | 0.18 | 0.16       | .873   | 37    | -0.190     | -0.29 | 9.9 × 10⁻⁵  |
|          |                          |         |                  |              |            |        |                  |        |                  |        |                  |        |
|          | Alcohol/week             | 40      | 37.41            | -0.01      | -0.04      | 0.02   | -0.02 | -0.03       | -0.03  | 0.01   | -0.02       | .545  | 0.34 | 0.04       | .341   | 35    | -0.020     | 0.04  | .043   |
|          |                          |         |                  |              |            |        |                  |        |                  |        |                  |        |
|          | Alcohol use disorder     | 46      | 37.00            | 0.26       | 0.08       | 0.44   | 0.16  | -0.10       | 0.42   | 0.14   | -0.12       | .288  | n.a.  | n.a.       | n.a.   | 43    | 0.310      | 0.10  | .003   |
|          |                          |         |                  |              |            |        |                  |        |                  |        |                  |        |
| Smoking  | Initiation DB            | 317     | 27.53            | 0.17       | 0.14       | 9.9    | 0.15  | 0.05       | 8.0    | 0.25   | 0.05       | .044  | n.a.  | n.a.       | n.a.   | 296   | 0.186      | 0.17  | 6.4 × 10⁻⁷² |
|          |                          |         |                  |              |            |        |                  |        |                  |        |                  |        |
| Alcohol/ | week DB                 | 80      | 28.61            | 0.04       | -0.13      | 0.21   | .632  | -0.12       | 0.22   | -0.17 | -2.07       | .859  | 0.13 | -0.24       | .06   | 74    | 0.0719     | 0.05  | 2.61   |
|          |                          |         |                  |              |            |        |                  |        |                  |        |                  |        |
| Alcohol  | Dis. 5 × 10⁻⁸            | 7       | 26.35            | -0.02      | -0.06      | 0.02   | -0.02 | -0.06       | 0.02   | -0.01 | -0.08       | .833  | n.a.  | n.a.       | n.a.   | 7     | 0.00016    | 0.04  | .993   |
|          |                          |         |                  |              |            |        |                  |        |                  |        |                  |        |
| Alcohol  | Dis. 1 × 10⁻⁵            | 25      | 21.94            | -0.01      | -0.03      | 0.01   | .328  | -0.01       | 0.03   | -0.01 | -0.06       | .654  | n.a.  | n.a.       | n.a.   | 24    | -0.0078    | 0.02  | .365   |

n.a. indicates that MR-Egger results were not reported because of limited reliability based on the I² measure being < 0.60. F > 10 generally indicates the instrument is sufficiently strong. DB, disruptive behavior; Dis., disorder; F, F-statistic indicating instrument strength; GSMR, generalized summary data-based Mendelian randomization; IVW, inverse variance weighted regression; SIMEX, simulation extrapolation; SNP, single nucleotide polymorphism.
**Differential Gene Expression.** To pinpoint relevant brain areas associated with DB and RTB more accurately, we proceeded with a local differential gene expression approach using stratified LDSC (31). For DB, we identified significant enrichment after Bonferroni correction exclusively in the dentate nucleus (cerebellum, \(Z = 3.35, \ p = .0004\)) (Figure 4A; Table S23 in Supplement 2). For RTB, there was also enrichment in the dentate nucleus (cerebellum, \(Z = 3.22, \ p = .0006\)) and additional enrichment in the cuneiform nucleus (brain stem, \(Z = 4.74, \ p = 1.06 \times 10^{-5}\)) (Figure 4B), claustrum (cortex, \(Z = 4.08, \ p = 2.27 \times 10^{-5}\)) (Figure 4C), paracentral lobule (cortex, \(Z = 3.35, \ p = 0.0004\)) (Figure 4D), lateral amygdaloid nucleus (subcortex, \(Z = 3.12, \ p = .0009\)) (Figure 4E), and preoptic region (subcortex, \(Z = 3.34, \ p = .0004\)) (Figure 4F; Table S24 in Supplement 2).

**Single-Cell Analysis.** Finally, we investigated the involvement of specific brain cell types using LDSC. For DB, this revealed enrichment in excitatory neurons of the prefrontal cortex (\(Z = 2.88, \ p = .002\)) and the hippocampal CA3 region (\(Z = 2.85, \ p = .002\)) (Table S25 in Supplement 2). For RTB, we found enrichment in excitatory prefrontal cortex neurons (\(Z = 3.14, \ p = .0008\)) (Table S26 in Supplement 2).

**DISCUSSION**

In a bottom-up approach to evaluate overlap in genetic and neurobiological backgrounds underlying externalizing symptoms, we collected 13 publicly available GWAS summary statistics for a range of externalizing phenotypes. Assessment of genetic resemblance indicated a categorization into two externalizing-related factors, one characterized by DB and another by RTB, that together explained 53% of the total genetic variance. Meta-analyzing the factor-specific phenotypes yielded 50 loci for DB and 257 loci for RTB.

Clustering of aggression, anger, and irritability items in the DB factor fits very well with the general finding that these are highly comorbid behaviors and also cluster together in diagnoses of conduct and oppositional defiant disorder (33).

Addition of ADHD aligns with the fact that DBs coexist most commonly with ADHD (36) and with evidence of strong genetic resemblance between ADHD and conduct and oppositional defiant disorder (37). The phenotypes of the RTB factor related to pursuing risk (21). Three RTB traits (general risk tolerance, number of sexual partners, lifetime cannabis use) similarly clustered together in another recent GWAMA (38). Addition of antisocial behavior aligns with the fact that antisocial behavior frequently includes thoughtlessness, self-centered, and immediately rewarding acts, which are also typical for risk taking. The present RTB phenotypes related to substance use, speeding, and sexual promiscuity are closely linked to social misconduct (38,39). In genomic structural equation modeling, antisocial behavior also loaded on DB, which is concordant with observations from the Hierarchical Taxonomy of Psychopathology consortium that antisocial behavior is a broad construct and also alludes to disruptive phenotypes (1,40).

Assessment of genetic relationships with other traits indicated associations for DB with higher neuroticism and lower conscientiousness and agreeableness, corroborating previous findings (41). For RTB, we found associations with higher extraversion, higher openness to experience, and lower conscientiousness, which is also in line with earlier reports (42).

Another notable result was that DB, much more than RTB, showed negative genetic associations with indices of mental and physical health. Previous studies indicated associations with mental and physical health problems for both disruptive-ADHD and risk-taking behaviors (43–46). This finding of stronger genetic correlations for DB points to the possible distinction that DB phenotypes, characterized by affective instability/irritability and impaired cognitive control, are generally central to mental and physical health problems (44,47), whereas for RTB, mental health problems are more often at the core (45), resulting in risk behaviors, such as substance use, that ultimately affect physical health (46).

In addition, we applied MR to explore causal relationships between the DB factor and smoking and alcohol use. In full agreement with previous findings (15,48), we obtained evidence that DB traits promote substance use (strong evidence...
for an increased odds of smoking initiation and decreased odds to successfully quit smoking and weak evidence for increased odds of alcohol use disorder). This causal direction is also consistent with expectations for behavior development, in that DB traits are more likely to occur at younger ages than risk behaviors. Smoking and alcohol use are not yet appropriate and less feasible during early childhood (15, 48). However, we also replicated the observation that smoking might causally increase the risk for DB traits. This finding could have important consequences for intervention strategies but needs further inquiry. To distinguish a causal influence from horizontal pleiotropy, i.e., when a genetic variant influences the two traits through independent pathways, dose-response effects should be included in future MR analyses (taking cigarettes per day as the exposure and applying stratification on smoking status).

Biological annotation by gene set analysis and stratified LDSC converged on genetic enrichment in the brain for both DB and RTB traits. An atlas of differential brain region gene expression pointed specifically to the dentate nucleus of the cerebellum, for both DB and RTB, and the cuneiform nucleus of the brainstem, preoptic region of the hypothalamus, lateral nucleus of the amygdala, and the claustrum and paracentral cortical lobule, which were significant for RTB.

Involvement of these brain regions confirms current hypotheses on the neurobiological background of externalizing symptoms but also points to the importance of relatively underrepresented regions. Especially for the medial part of the preoptic hypothalamic region, animal studies indicate an important role in stimulating aggression to facilitate reproduction (49), as well as in modulating mesolimbic activity involved in reward processing (50). The contribution of the amygdala to emotional processing is also generally recognized (51), and there is ample evidence that amygdala–frontal cortex network abnormalities predispose to externalizing behaviors (14), including aggression and risk taking (52, 53). In addition, amygdala volume reductions are commonly found in ADHD (54–56). Furthermore, a recent study showed that GABAergic neurons in the medial amygdala project to the medial preoptic area to regulate reward from social stimuli by controlling the release of dopamine in the nucleus accumbens (57).

The other brain regions have been implicated less frequently, despite previous evidence for associations. For the cerebellar dentate, a relevant role in the development of reinforcement learning relevant to addiction has been indicated (58), and there is evidence that abnormal development of corticocerebellar connections contributes to ADHD (59). The cuneiform nucleus is part of the mesencephalic locomotor control network (60) and related to autonomic fear and stress responses (61, 62). The claustrum is richly interconnected with almost all regions of the cerebral cortex and hypothesized to play a role in regulating attention and resilience to distraction (63, 64), which links this region to ADHD (65). The anterior section of the paracentral lobule includes motor control regions but has also been linked to executive control function and attention orienting (22, 23). Single-cell analysis further elucidated the important role of the CNS for both DB and RTB and pointed to glutamatergic prefrontal neurons for DB and RTB and pyramidal hippocampal neurons for DB. Glutamatergic neurotransmission has been associated with alcohol and tobacco use (33) and RTB (21) before, but more research is needed to address the role of these specific cell types and their functions on the externalizing phenotypes.

Finally, testing for specific biological mechanisms using gene set analysis indicated associations for gene sets primarily related to neuronal development for both DB and RTB. In addition, synaptic functions and transcription regulation were identified for RTB. These three biological mechanisms have been related to a wide range of psychiatric disorders (66) and may play a broad role in behaviors and brain-related traits.

This study comes with a number of limitations. First, the two identified factors DB and RTB combine several distinct phenotypic measures, which inevitably leads to etiologic heterogeneity (reflected by the imperfect genetic correlations). In addition, phenotyping could be as poor as a single question (e.g., about regular use of a substance). Poor and heterogeneous phenotyping, with the trade-off of having large sample sizes available, are a common phenomenon in GWAS meta-analyses of single traits (67) or multiple traits (22) and generally results in lower SNP heritability estimates compared with the individual traits, which was also the case here for DB (4%) and RTB (2%). Second, we incorporated three measures on irritability, which seems redundant. However, the mutual genetic correlation between these phenotypes was not perfect ($r_g < 1$), and therefore, we note that including GWAS statistics of all three phenotypes does provide additional genetic information in GWAMAs. Third, in the MR analyses, the tested effects from smoking and alcohol use to DB did not fully comply with the expected temporal order. The aggression and ADHD variables of DB included children (which amounts to ~5% of the total number of DB participants). This means that smoking and alcohol use were, for those variables, not valid exposures. In addition, we limited assessment of causal directions to the relationship between the DB factor and smoking and alcohol use. We also demonstrated interesting relationships of DB and RTB with several other traits, such as personality and health characteristics. Given the complexity of the included study cohorts forming the different summary statistics, it was beyond the scope of this study to also test for causal relationships with these traits. Fourth, except for ADHD, the GWAS data in our study related to behavioral extremes that did not exceed the diagnostic threshold. Given previous evidence of genetic heterogeneity between nondiagnosed and diagnosed individuals (68, 69), we emphasize that the findings in this study apply in particular to general rather than excessive diagnosed externalizing behaviors. Finally, regarding the biological annotation analyses, we must consider limitations in drawing conclusions about different findings for DB and RTB. DB and RTB are moderately correlated ($r_g = 0.33$) and therefore not completely independent. In addition, the effective sample size of RTB is ~3 times larger than that of DB (1,506,537 vs. 523,150), so there is likely more power to detect significant associations for RTB. For example, differential gene expression indicated more brain regions with statistically significant enrichment for RTB, but we note that the same regions also showed relatively increased enrichment among the 211 brain regions tested for DB.

In summary, we extend previous findings involving externalizing behavior and provide further evidence for common genetic architectures, particularly for different DBs and RTBs.
Follow-up evaluations of the data in this study indicated genetic relationships with personality traits and mental and physical health behaviors in agreement with the DB versus RTB categorization and highlighted possible bidirectional causal relationships between DB and substance use traits. Biological annotation revealed generally similar gene-brain mappings that mediate the predisposition to comorbid externalizing phenotypes. Possible subtle differences in neurobiological backgrounds between DB and RTB traits may be resolved in future studies based on higher-resolution genetic association and gene expression data that are becoming increasingly available.

ACKNOWLEDGMENTS AND DISCLOSURES
We thank the research participants and employees of 23andMe and UK Biobank for making this work possible. In addition, we thank the Social Science Genetic Consortium, the EAGLE consortium, and the International Cannabis Consortium for making their summary statistics publicly available. Without their data sharing policy, this study was not possible.

This study used genome-wide association study summary statistics published by 23andMe and the UK Biobank. All input data used in the present study are from public sources (listed in Table 1). Output data are accessible published by 23andMe and the UK Biobank. All input data used in the present study are from public sources (listed in Table 1). Output data are accessible.

REFERENCES
1. Krueger RF, Kotov R, Watson D, Forbes MK, Eaton NR, Ruggiero CJ, et al. (2018): Progress in achieving quantitative classification of psychopathology. World Psychiatry 17:292–293.
2. Krueger RF, South SC (2009): Externalizing disorders: Cluster 5 of the DSM-V and ICD-11. Psychiatr Med 39:2061–2070.
3. Beauchaine TP, Zisner AR, Sauder CL (2017): Trait impulsivity and the externalizing spectrum. Annu Rev Clin Psychol 13:343–368.
4. Ahmad SI, Hinshaw SP (2017): Attention-deficit/hyperactivity disorder, trait impulsivity, and externalizing behavior in a longitudinal sample. J Abnorm Child Psychol 45:1077–1089.
5. GBD 2016 Alcohol Collaborators (2018): Alcohol use and burden for 195 countries and territories, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. Lancet 392:1015–1035.
6. Bartels M, Hendriks A, Mauri M, Kraoehl E, Whipp A, Bohus K, et al. (2018): Childhood aggression and the co-occurrence of behavioural and emotional problems: Results across ages 3–16 years from multiple raters in six cohorts in the EU-ACTION project. Eur Child Adolesc Psychiatry 27:1105–1121.
7. Mullins-Sweatt SN, DeShong HL, Lengel GJ, Helle AC, Krueger RF (2019): Disinhibition as a unifying construct in understanding how personality dispositions undergird psychopathology. J Res Pers 80:55–61.
8. Jang KL, Vernon PA, Livesley WJ (2000): Personality disorder traits, family environment, and alcohol misuse: A multivariate behavioural genetic analysis. Addiction 95:873–888.
9. Krueger RF, Hicks BM, Patrick CJ, Carlson SR, Iacono WG, McGue M (2002): Etiologic connections among substance dependence, antisocial behavior, and personality: Modeling the externalizing spectrum. J Abnorm Psychol 111:411–424.
10. Venables NC, Hicks BM, Yancey JR, Kramer MD, Nelson LD, Strickland CM, et al. (2017): Evidence of a prominent genetic basis for associations between psychoneurotic traits and common mental disorders. Int J Psychophysiol 115:4–12.
11. Young SE, Friedman NP, Miyake A, Willcut EG, Corley RP, Haberstick BC, Hewitt JK (2009): Behavioral disinhibition: Liability for externalizing spectrum disorders and its genetic and environmental role in response inhibition across adolescence. J Abnorm Psychol 118:117–130.
12. Noordermeer SDS, Luman M, Oosterlaan J (2016): A systematic review and meta-analysis of neuroimaging in oppositional defiant disorder (ODD) and conduct disorder (CD) taking attention-deficit hyperactivity disorder (ADHD) into account. Neuropsychol Rev 26:44–72.
13. Yang Y, Raine A (2009): Prefrontal structural and functional brain imaging findings in antisocial, violent, and psychopathic individuals: A meta-analysis. Psychiatry Res 174:81–88.
14. Ameis SH, Ducharme S, Albaugh MD, Hudziak JJ, Botteron KN, Lepage C, et al. (2014): Cortical thickness, cortico-amygdalar networks, and externalizing behaviors in healthy children. Biol Psychiatry 75:65–72.
15. Treur JL, Demontis D, Smith GD, Silallis H, Richardson TG, Wiers RW, et al. (2021): Investigating causality between liability to ADHD and substance use, and liability to substance use and ADHD risk, using Mendelian randomization. Addict Biol 26:e12849.
16. Pasman JA, Verweij KJH, Gerring Z, Stringer S, Sanchez-Roige S, Treur JL, et al. (2018): GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia. Nat Neurosci 21:1161–1170.
17. Tielbeek JJ, Johansson A, Polderman TJ, Rautiainen MR, Jansen P, Taylor M, et al. (2017): Genome-wide association studies of a broad spectrum of antisocial behavior. JAMA Psychiatry 74:1242–1250.
18. Pappa I, St Pourcain B, Benke K, Cavadino A, Hakulinen C, Nivard MG, et al. (2018): A genome-wise approach to children’s aggressive behavior: The EAGLE consortium. Am J Med Genet B Neuropsychiatr Genet 171:562–572.
19. Suddow C, Gallagher J, Allen N, Beral V, Burton P, Danesh J, et al. (2015): UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 12:e1001779.
20. Demontis D, Walters RK, Martin J, Mattei Haff M, Als TD, Agerbo E, et al. (2019): Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Genet 51:63–75.
21. Karlsson Lirim R, Birol P, Kong E, Meddins SFW, Wedow R, Fontana MA, et al. (2019): Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nat Genet 51:245–257.
22. Baselinas BML, Jansen R, Ip HF, van Dongen J, Abdellaoui A, van de Weijer MP, et al. (2019): Multivariate genome-wide analyses of the well-being spectrum. Nat Genet 51:445–451.
23. Gravel S (2012): Population genetics models of local ancestry. Genetics 191:607–619.
24. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics Consortium, et al. (2015): LD Score regression distinguishes confounding from
polygenicity in genome-wide association studies. Nat Genet 47:291–295.
25. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. (2015): An atlas of genetic correlations across human diseases and traits. Nat Genet 47:1236–1241.
26. Ward JH Jr (1963): Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58:236–244.
27. Calinski T, Harabasz J (1974): A dendrite method for cluster analysis. Commun Stat 3:1–27.
28. Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, et al. (2019): Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. Nat Hum Behav 3:513–525.
29. De Leeuw CA, Mooij JM, Heskes T, Posthuma D (2015): MAGMA: Generalized gene-set analysis of GWAS data. PLoS Comput Biol 11: e1004219.
30. GTEx Consortium (2015): Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: Multisite tissue gene regulation in humans. Science 348:648–660.
31. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, et al. (2012): An anatomically comprehensive atlas of the adult human brain transcriptome. Nature 499:391–399.
32. Habib N, Avraham-David I, Basu A, Burks T, Shekhar K, Hofree M, et al. (2017): Massively parallel single-nucleus RNA-seq with DroNc-seq. Nat Methods 14:955–960.
33. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. (2019): Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. Nat Neurosci 21:1656–1668.
34. American Psychiatric Association (2013): Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition Washington, DC: American Psychiatric Association Publishing.
35. Kunwar A, Dewan M, Faraone SV (2007): Treating common psychiatric disorders associated with attention-deficit/hyperactivity disorder. Expert Opin Pharmacother 8:555–562.
36. Kim H, Viken RJ, Kaprio J, Pulkkinen L, Rose RJ (2005): Association studies of ADHD symptoms: Links between personality traits and clinical symptoms. J Abnorm Child Psychol 33:219–229.
37. Karlsson Linner R, Mallard TT, Barr PB, Sanchez-Roige S, Madole JW, Driver MN, et al. (2021): Multivariate genomic analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. Nat Neurosci 24:831–842.
38. Moulton EA, Elman I, Becerra LR, Goldstein RZ, Borsook D (2014): The amygdala in adolescents with attention-deficit/hyperactivity disorder: Structural and functional correlates of delay aversion. World J Biol Psychiatry 15:673–684.
39. Hoogman M, Braaten J, Hibr DP, Mennes M, Zivs MP, et al. (2017): Subcortical brain volume differences in participants with attention deficit hyperactivity disorder and adults: A cross-sectional mega-analysis [published correction appears in Lancet Psychiatry 2017; 4:436]. Lancet Psychiatry 4:310–319.
40. Hu RK, Zuo Y, Li T, Wang J, Meera P, Wu YE, Hong W (2021): An amygdala-to-hypothalamus circuit for social reward. Nat Neurosci 24:831–842.
41. Moulton EA, Elman I, Becerra LR, Goldstein RZ, Borsook D (2014): The cerebellum and addiction: Insights gained from neuroimaging research. Addict Biol 19:317–331.
42. Shaw P, Iishi-Takahashi A, Park MT, Devenyi GA, Zibman C, Gaskierek S, et al. (2018): A multicohort, longitudinal study of cerebellar development in attention deficit hyperactivity disorder. J Child Psychol Psychiatry 59:1114–1123.
43. Caggiano V, Lerais R, Goﬁ-Enro H, Masini D, Bellardita C, Bouvier J, et al. (2018): Midbrain circuits that set locomotor speed and gait selection. Nature 553:455–460.
44. Silva BA, Gross CT, Grätz J (2016): The neural circuits of innate fear: Detection, integration, action, and memorization. Learn Mem 23:544–555.
45. Korte SM, Jaarsma D, Luiten PG, Bohus B (1992): Mesencephalic cuneiform nucleus and its ascending and descending projections serve stress-related cardiovascular responses in the rat. J Auton Nerv Syst 41:157–176.
46. Goll Y, Attan G, Citri A (2015): Attention: The claustrum. Trends Neurosci 38:486–495.
47. Attan G, Terem A, Perez-Rivlin N, Sehrawat K, Gonzales BJ, Pozner G, et al. (2018): The claustrum supports resilience to distraction. Curr Biol 28:2752–2762.e7.
65. Dickstein SG, Bannon K, Castellanos FX, Milham MP (2006): The neural correlates of attention deficit hyperactivity disorder: An ALE meta-analysis. J Child Psychol Psychiatry 47:1051–1062.

66. Cross-Disorder Group of the Psychiatric Genomics Consortium (2019): Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. Cell 179:1469–1482.e11.

67. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abellaou A, et al. (2018): Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet 50:668–681.

68. Johnson EC, Demontis D, Thorgeirsson TE, Walters RK, Polimanti R, Hatoum AS, et al. (2020): A large-scale genome-wide association study meta-analysis of cannabis use disorder. Lancet Psychiatry 7:1032–1045.

69. Kranzler HR, Zhou H, Kember RL, Vickers Smith R, Justice AC, Damrauer S, et al. (2019): Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations [published correction appears in Nat Commun 2019; 10:2275 and Nat Commun 2019; 10:4050]. Nat Commun 10:1499.