Published in final edited form as:

J Clin Child Adolesc Psychol. 2017 ; 46(6): 824–839. doi:10.1080/15374416.2015.1070353.

Dimensionality and Genetic Correlates of Problem Behavior in Low-income African American Adolescents

Shawn J. Latendresse,

Baylor University, Waco, TX

David B. Henry,

University of Illinois at Chicago, Chicago, IL

Steven H. Aggen,

Virginia Commonwealth University, Richmond, Va

Gayle R. Byck,

Northwestern University, Chicago, IL

Alan W. Ashbeck,

Northwestern University, Chicago, IL

John M. Bolland,

University of Alabama at Birmingham, Birmingham, AL

Cuie Sun,

Virginia Commonwealth University, Richmond, VA

Brien P. Riley,

Virginia Commonwealth University, Richmond, VA

Brian Mustanski^{*}, and

Northwestern University, Chicago, IL

Danielle M. Dick^{*}

Virginia Commonwealth University, Richmond, VA

Abstract

Objective—Researchers have long observed that problem behaviors tend to cluster together, particularly among adolescents. Epidemiological studies have suggested that this covariation is due, in part, to common genetic influences, and a number of plausible candidates have emerged as targets for investigation. To date, however, genetic association studies of these behaviors have mostly focused on unidimensional models of individual phenotypes within European American samples.

Correspondence should be addressed to Shawn J. Latendresse, Department of Psychology & Neuroscience, Baylor University, One Bear Place #97334, Waco, TX 76798 (shawn_latendresse@baylor.edu); and Brian Mustanski, Department of Medical Social Sciences, Northwestern University, 625 N. Michigan Avenue, Suite 2700, Chicago, IL 60611 (brian@northwestern.edu).

^{*}Denotes equal contributions as senior author for this manuscript. All content is the sole responsibility of the authors.

Method—Herein, we compared a series of confirmatory factor models to best characterize the structure of problem behavior (*alcohol and marijuana use, sexual behavior, and disruptive behavior*) within a representative community-based sample of 592 low-income African American adolescents (50.3% female), aged 13 to 18. We further explored the extent to which three genes previously implicated for their role in similar behavioral dimensions (*CHRM2, GABRA2, and OPRM1*) independently accounted for variance within factors specified in the best fitting model. Supplementary analyses were conducted to derive comparative estimates for the predictive utility of these genes in more traditional unidimensional models.

Results—Findings provide initial evidence for a bifactor structure of problem behavior among African American adolescents, and highlight novel genetic correlates of specific behavioral dimensions otherwise undetected in an orthogonal syndromal factor.

Conclusions—Implications of this approach include increased precision in the assessment of problem behavior, with corresponding increases in the reliability and validity of identified genetic associations. As a corollary, the comparison of primary and supplementary association analyses illustrates the potential for overlooking and/or overinterpreting meaningful genetic effects when failing to adequately account for phenotypic complexity.

Keywords

Substance Use; Sexual behavior; Disruptive Behavior; Molecular Genetics; Ethnic minority youth

By the end of high school, nearly four out of five adolescents in the U.S. have consumed alcohol, almost half have used marijuana, and nearly one quarter have had four or more sexual partners (Centers for Disease Control, 2012). In addition, up to 25% of adolescents in grades seven through twelve have reported recently engaging in some type of disruptive behavior (e.g., shoplifting, skipping school, and damaging property; Bartlett, Holditch-Davis, & Belyea, 2007). Although these specific risk-taking behaviors are all somewhat developmentally normative, the relative pervasiveness is still quite alarming, particularly when considering that the potential long-term consequences to self and society include substance use disorders, antisocial behaviors leading to incarceration or injury, sexually transmitted diseases, and unwanted pregnancies.

Perhaps even more concerning is the frequency with which these behaviors tend to co-occur in adolescents (McGee & Newcomb, 1992), and the increased likelihood of risk that accompanies each additional occurrence (McGue & Iacono, 2005). These issues have received a great deal of attention in the nearly four decades since Jessor & Jessor (1977) initially posited that frequently observed co-occurrences among a set of *problem behaviors* (illicit drug use, problem drinking, delinquent behavior, and precocious sexual intercourse; Donovan & Jessor, 1985), reflects an underlying *problem behavior syndrome* (PBS). Findings from a recent latent class analysis of health risk behaviors within a representative sample of early-adolescents illustrates the continued prevalence of this phenomenon, suggesting that one in four exhibits a behavioral profile characterized by high probabilities of substance use, delinquency, and precocious sexual behavior (Hair, Park, Ling, & Moore, 2009). Moreover, another recent study suggests that once PBS is manifest, one's ability to

desist from specific behaviors becomes increasingly more difficult (Monahan, Rhew, Hawkins, & Brown, 2013).

Notably, Jessor's operationalization of PBS as problematic, socially undesirable or unconventional behaviors that cause concern and/or elicit social control responses (1977) is generally synonymous with what others have referred to as externalizing behavior or externalizing symptomatology (Seglem, Torgersen, Ask, & Waaktaar, 2015). All of these are broadband constructs reflecting shared variance across a set of more narrowly defined, but threshold free behavioral continua. In contrast, both externalizing psychopathology and externalizing spectra are generally operationalized as overarching constructs consisting either exclusively of disordered behaviors, or some combination of clinically derived disorders and symptoms (Tackett, 2010). Moreover, although models of PBS have included various constellations of behavior over the years, the four original dimensions have remained constant (e.g., see Donovan, Jessor, & Costa, 1991).

The empirical validation of PBS has generally relied on confirmatory factor models in which composite scores for the distinct domains of problem behavior simultaneously load onto a single common factor. As such, the factor reflected shared variability across a broad range of externalizing behaviors, and its structure has been replicated in adolescents and young adults across both U.S. (Donovan & Jessor, 1985; Donovan, Jessor, & Costa, 1988) and cross-national samples (Vazsonyi et al., 2010). Other studies utilizing multiple indicators of domain-specific externalizing problems in place of summary scores have suggested that a higher order structure significantly improves model fit (McGee & Newcomb, 1992; Newcomb et al., 2002). That is, at the item level, behavioral indicators are more highly correlated within versus between domains, so those inter-item relationships are captured via domain-specific first order confirmatory factors, and covariance among the first order factors is accounted for by a second order factor representing the syndromal nature of problem behavior. While much of the hierarchical work on problem behavior has been conducted across diverse samples, this higher order factor structure has also been validated within a sample of African American (AA) adolescents (Resnicow, Ross-Gaddy, & Vaughan, 1995).

Overall, the literature implies that beyond accounting for the shared variance across forms of problem behavior, domain specific factors can also yield important information, suggesting the need to explore both syndromal and unique aspects (Allen, Leadbeater, & Aber, 1994). Toward this end, a more comprehensive and systematic evaluation of the underlying factor structure may help clarify the etiology of problem behavior (Reise, Morizot, & Hays, 2007), as our present understanding is largely contingent upon how these constructs have been operationally organized.

Though principally rooted in a multifaceted system of psychosocial factors (Jessor, Van Den Bos, Vanderryn, Costa, & Turbin, 1995), Jessor and colleagues explicitly denote genetics as an important explanatory domain with respect to problem behavior (Jessor, 1987; Jessor, Donovan, & Costa, 1991). Consistent with this perspective, twin and family studies indicate shared sources of genetic liability across a wide range of externalizing problems/disorders. At the latent level, variance decomposition analyses in twin samples provide evidence for common genetic etiologies with respect to alcohol dependence (AD) and various disruptive

behavior disorders, including oppositional defiant disorder (ODD), conduct disorder (CD), and antisocial personality disorder (ASPD) (Kendler, Prescott, Myers, & Neale, 2003; Markon & Krueger, 2005), for alcohol and marijuana use (Sartor et al., 2010), and across adolescent conduct problems and risky sexual behavior (Verweij, Zietsch, Bailey, & Martin, 2009). Moreover, genome-wide linkage analyses intimate regions of shared liability for substance dependence (SD) and ASPD (Stallings et al., 2005), and for AD and CD (Dick et al., 2004).

In the present era of large-scale gene finding efforts, genome-wide association studies (GWAS) have provided increasingly detailed information, allowing us to more precisely identify specific candidates for association, from whole genes down to the level of single nucleotide polymorphisms (SNPs). As a result, genes initially identified as candidates for specific disordered behaviors, such as AD (*CHRM2*, *GABRA2*) and opioid dependence (*OPRM1*), have subsequently been implicated as playing a broader role in relation to discrete non-disordered forms of substance use (Corley et al., 2008; Dick et al., 2008; Edenberg et al., 2004; Ehlers, Lind, & Wilhelmsen, 2008; Ittiwut et al., 2012; Kendler et al., 2011; Lind et al., 2008; Luo, Kranzler, Zhao, & Gelernter, 2003; Smelsen et al., 2012; Zhang et al., 2006), disruptive behavior (Dick et al., 2006; Dick et al., 2008), and inhibitory control (Dick et al., 2008; Hendershot, Bryan, Feldstein Ewing, Claus, & Hutchison, 2011), as well as generalized externalizing behavior (Dick, 2007; Dick et al., 2008; Dick et al., 2009; Latendresse et al., 2011).

Unfortunately, despite mounting evidence suggesting (1) shared genetic variance across many common externalizing constructs, and (2) improved measurement and increased validity of externalizing phenotypes and resulting associations when explicitly modeling known multidimensionality, all attempts to identify specific genetic variants associated with externalizing problems/disorders to date appear to rely on independent analyses of unitary phenotypes derived to reflect unidimensional disorders, behaviors, or broad theoretical constructs. Therefore, any meaningful covariation amongst related externalizing dimensions might have been misallocated in extant studies of association between genotypes and unconditional externalizing phenotypes.

In addition, genetic association studies are often conducted within racially homogeneous subgroups. Thus, many studies test solely for associations within the largest group in their sample, typically European Americans (EA) within U.S. samples, or simply include race (a frequently used proxy for differentiating discrete ancestral populations; Risch, Burchard, Ziv, & Tang, 2002) as a covariate in pooled samples analyses (cf. Ittiwut et al., 2012; Kendler et al., 2011; Smelsen et al., 2012). As such, there has been little research specifically exploring genetic associations among AA youths, despite the fact that they experience a number of disparities related to the consequences of problem behaviors.

Herein, we have compared a series of confirmatory factor models characterizing the structure of problem behavior within a representative community-based sample of low-income AA adolescents. The relative value of a superior phenotypic measurement model was assessed by separately regressing factor scores derived from the best fitting model and traditional unidimensional confirmatory factor models commonly used in genetic association

studies onto a common set of SNPs. The present study focused on variants spanning three genes previously implicated in association with general and/or specific dimensions of externalizing behavior/disorder: *CHRM2*, *GABRA2*, and *OPRM1*. Lastly, converging/diverging results across distinct models, and broader implications for future genetic association studies with complex behaviors were discussed.

Method

Participants and Procedures

Sample selection and recruitment—The Gene, Environment, Neighborhood Initiative (GENI) constitutes a representative community sample of 592 adolescents aged 13 through 18 and their primary caregiver (see Mustanski et al., 2013). GENI participants were recruited from the Mobile Youth Study (MYS), a community-based, multiple cohort longitudinal study with annual data collection. MYS has been described in detail elsewhere (Bolland, 2007; Park, Lee, Bolland, Vazsonyi, & Sun, 2008) and consists of adolescents and caregivers from predominantly AA, very low-income inner-city neighborhoods in the Mobile, Alabama metropolitan statistical area (United States Census Bureau, 2002: e.g., > 92% AA, > 27% poverty rate among youth). Previous comparisons between MYS participants who did and did not participate in the GENI study suggest no significant differences with respect to risk behaviors, and only trivial demographic differences (Byck, Bolland, Dick, Ashbeck, & Mustanski, 2013).

Participant interviews—The Institutional Review Boards at the University of Alabama, Northwestern University, and Virginia Commonwealth University approved procedures for this study. Participation in GENI involved an approximately two and a half hour interview for both the adolescent and his/her caregiver. Interviews were conducted between March 2009 and October 2011. Written parental consent and youth assent were obtained. Caregivers and adolescents were compensated for participation.

Biological data collection—DNA was obtained via saliva sample using Oragene collection kits, under the supervision of a specially trained interviewer. Samples were coded anonymously and mailed to the laboratory of Dr. Brien Riley at the Virginia Institute for Psychiatric and Behavioral Genetics (VIPBG) at Virginia Commonwealth University, where DNA extraction was performed. In total, DNA samples were obtained from 573 adolescents (50.3% female), representing 96.8% of the total GENI sample. The racial composition within this group was derived via interviewer administered Family Demographic Caregiver Reports, wherein each adolescent was characterized as a member of one of six discrete “Race/Ethnicity” groups: *Black* (N=566), *White* (N=2), *Mixed Race* (N=5), *Asian* (N=0), *Hispanic/Latino* (N=0), and *Other* (N=0).

Genotyping—From these samples, a total of 69 single nucleotide polymorphisms (SNPs) were genotyped across *CHRM2* (25), *GABRA2* (20), and *OPRM1* (24) (see Table 1). Genotyping was conducted at the VIPBG using fluorescence polarization detection of template-directed dye-terminator incorporation with appropriate AcycloPrime SNP detection kit for specific polymorphisms (PerkinElmer, Boston) and an automated allele-scoring

platform (Van den Oord, Jiang, Riley, Kendler, & Chen, 2003). First, sixteen SNPs previously typed in EA samples by our group were force-included in the tagging set for the purpose of future comparison. Next, TaqMan assays were not available for three previously typed SNPs, so SNPs in complete linkage disequilibrium (LD) with those EA typed SNPs (i.e., $R^2 = 1.0$) were selected as proxies. Finally, to complete LD tagging across each of these previously identified regions of genetic risk for externalizing phenotypes, a supplementary set of 50 SNPs was selected from among those identified in the Nigerian Yoruba population of the HapMap project, thus capturing any additional genetic variability existing within individuals of more recent African descent.

Genotyping success rates for SNPs in these three genes ranged from 98.5–98.9%, and duplicate genotyping produced concordance rates of 100%. By way of quality control, all genetic data were subjected to a set of inclusionary thresholds. First, individual DNA samples yielding a gene-wise genotyping success rate of less than 80% were deemed unreliable and removed from consideration for inclusion in the analytic dataset. For all remaining data, any individual SNP with a sample-wise genotyping success rate of less than 80% was also excluded from the analytic dataset. Of the 573 participants from whom DNA data were collected, less than 1% of individuals failed to meet the first threshold for any given gene. Moreover, none of the markers genotyped were excluded on the basis of the second criteria. Finally, no SNPs were eliminated from analyses for significantly deviating from within-race calculations for Hardy-Weinberg equilibrium ($p < .001$). All SNP chromosomal positions and allele identities are depicted with respect to the genomic (+) strand.

Both allelic frequencies and LD structure have been shown to differ substantially between populations of discrete ancestral descent (Lonjou et al., 2003). To guard against bias resulting from population stratification we limited all genetic analyses to the subsets of caregiver-reported “black” adolescents surviving quality control procedures for each gene: N 's = 561 (*CHRM2*) and 562 (*GABRA2*, *OPRM1*). In addition, reported race was validated via principal components analysis using the Admixture software package (Alexander, Novembre, & Lange, 2009), where the average estimated European ancestry among “black” adolescents was less than 20%, and none reached 50%. Haploview (Barrett, Fry, Maller, & Daly, 2005) was used to estimate LD across the full set of genotyped SNPs within each gene (Figure 1). This inter-SNP correlation reflects the degree to which analyses with individual markers represent unique versus non-independent tests of association. SNPSpD (Nyholt, 2004), a web-based test, was used to derive Bonferroni-like correction factors for multiple testing that account for the number of SNPs genotyped and the LD structure between them. In turn, the effective numbers of independent marker loci for our analyses were estimated to be 15 in *CHRM2* and *OPRM1*, and 10 in *GABRA2*, yielding adjusted significance thresholds (α/n) of .003 and .005, respectively. Based on these critical values, we had adequate statistical power (i.e., .80) to detect genetic main effects in traditional unidimensional factor models of size $R^2 = .02-.025$. Moreover, the multidimensional measurement model employed in the present study had the potential to increase this power to detect genetic associations by 20–99% (van der Sluis, Verhage, Posthuma, & Dolan, 2010).

Measures

Alcohol use—Use of alcohol over the past 12 months was assessed via four self-report items. First, a single dichotomous item inquiring about whether participants had consumed six or more drinks during that period (0 = *no*, 1 = *yes*). This item was taken from the Diagnostic Interview Schedule for Children version 4.0 (C-DISC; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2008), a widely used instrument for assessing psychiatric diagnoses among adolescents that is administered as a computerized, structured interview by trained lay interviewers (for details of administration, see Mustanski et al., 2013). In addition, three ordinal response items from the AIDS-Risk Behavior Assessment (ARBA; Donenberg, Emerson, Bryant, Wilson, & Weber-Shiffrin, 2001) were used to assess the frequencies of use and heavy drinking in the last 12 months (0 = *zero days*, 1 = *one or two days*, 2 = *monthly or less*, 3 = *two to three days per month*, 4 = *one or two days a week*, 5 = *three to five days per week*, 6 = *every day or almost every day*), as well as the average quantity consumed (0 = *zero drinks*, 1 = *one drink*, 2 = *two drinks*, 3 = *three drinks*, 4 = *four drinks*, 5 = *five drinks*, 6 = *six or more drinks*).

Marijuana use—Use of marijuana during the past 12 months was assessed via two dichotomous items (0 = *no*, 1 = *yes*) from the C-DISC: “*used*”, and “*used six or more times*”. In addition, overall frequency of marijuana use within the past year was assessed via self-report using a single item from the ARBA. Responses for this item were coded as follows: 0 = *zero times*, 1 = *once or twice*, 2 = *monthly or less*, 3 = *two to three times per month*, 4 = *one or two times a week*, 5 = *three to five times per week*, 6 = *every day or almost every day*.

Sexual behavior—Lifetime sexual behavior was assessed via five self-reported items from the HIV-Risk Assessment for Sexual Partnerships (H-RASP; Mustanski, Starks, & Newcomb, 2012), a computerized self-administered interview that was adapted from the ARBA to better assess sexual behavior and associated situational and contextual variables at the level of the sexual partnership. First, three items inquired about whether participants had ever engaged in oral, vaginal, and anal sex (0 = *no*, 1 = *yes*). Two additional items asked about the number of lifetime vaginal/anal sex partners, and the number of those partners with whom no condom was used during sex. Both of these items used an open response format, and participants who indicated abstinence on initial inquiries regarding vaginal and anal sex were coded zero.

Disruptive behavior—The Youth Self Report (YSR; Achenbach, 1991) was used to assess disruptive behaviors. This battery consists of 112 items, for which participants indicate whether each is “*not true*”, “*somewhat or sometimes true*”, or “*very or often true*” (scored 0, 1, and 2, respectively) with respect to their behavior *over the last six months*. Of these items, five are considered to reflect a DSM-like oppositional defiant problems subscale (e.g., “*I disobey my parents*”), and another 15 items are considered to reflect a DSM-like conduct problems subscale (e.g., “*I cut classes or skip school*”).

Analytic Plan

A series of unconditional confirmatory factor models (see Figure 2) were fit in *Mplus* (Muthén & Muthén, 1998–2010) using the weighted least squares mean and variance adjusted estimation procedure. One-factor, four-factor, second-order factor, and bifactor models were estimated sequentially, with the four-factor model estimated both with and without covariances between the four factors reflecting problem-specific dimensions. Since the standard chi-square fit statistic (χ^2) is sensitive to sample size (Bentler & Bonnet 1980), a relative chi-square ratio (χ^2/df ; Wheaton, Muthén, Alwin, & Summers, 1977) and several additional omnibus fit indices were considered when assessing overall model fit, including both the Tucker-Lewis index (TLI; Tucker & Lewis, 1973) and root-mean-square error of approximation (RMSEA; Browne & Cudeck, 1993). In general, larger TLI values and smaller relative chi-square ratios and RMSEA values indicate better fitting models. Currently adopted conventions suggest using a relative chi-square ratio smaller than 2.0 (Tabachnick & Fidell, 2007), TLI values .95, and RMSEA values .06 denote good fit (Hu & Bentler, 1999).

In the set of primary association analyses, partial main effects for the three candidate genes were assessed by simultaneously regressing all latent factors within the best-fitting model onto each measured genotype, while controlling for sex and age. Subsequently, in the supplementary association analyses, unidimensional confirmatory factor structures were separately modeled for all five of the aforementioned general and specific dimensions of problem behavior. As with the latent bifactor structure in the primary analyses, each of these supplementary phenotypes was independently regressed on SNPs from the three candidate genes in order to assess the extent to which disparate models of phenotype robustly or differentially determine the observed genetic associations.

Within the context of the social sciences, some additional explanation regarding the analytic strategy may be warranted. To start, we elected to test a seemingly large number of SNPs within each gene of interest relative to other candidate gene studies in the social science literature. The choice to do so reflects current practices in genetics (Dick, Latendresse, & Riley, 2011; Tabor, Risch, & Myers, 2002), and is based on issues often overlooked by social scientists. First, robust association with a given phenotype does not imply identification of a functional genetic variant, and even when evidence for the functional relevance of specific genetic loci exists within a given population, it cannot be assumed to generalize across ancestral lines. Moreover, a gene can include multiple causal variants, each contributing differentially to its functioning and the manner in which it influences susceptibility to the phenotype. Therefore, absent knowledge of all causal variants for a specific population within the genetic region(s) of interest, association analyses require enough markers to account for the observed LD structure of ancestral populations most similar to the sample being studied. Although this entails testing, and thus correcting for large numbers of variants, and increasingly so within more genetically heterogeneous samples (e.g., 25 *CHRM2* SNPs in the present sample, compared with 9 SNPs genotyped in a commensurate sample of EA adolescents; Latendresse et al., 2011), there is no justifiable shortcut if the goal is to identify potentially meaningful phenotypic influences within sparsely characterized regions of the genome.

Results

Confirmatory Factor Modeling

One-factor model—The first model posited that the covariance among item responses for all 32 indicators of problem behavior (Figure 2, Model A) can be best accounted for by a single underlying common factor. Goodness of fit for this unidimensional model was poor (fit statistics for all confirmatory factor models are summarized in Table 2), as indicated by a large relative chi-square ratio, sub-threshold TLI and RMSEA values ($\chi^2/df = 5.16$, TLI = .85, RMSEA = .08).

Four-factor models—The second model posited that the covariance among item responses for the 32 indicators would be best explained by allowing each of the four dimension-specific groups of indicators (i.e., alcohol use items, marijuana use items, sexual behavior items, and disruptive behavior items) to load onto independent (orthogonal) dimension-specific factors (Figure 2, Model B). Goodness of fit for this uncorrelated multidimensional model was also poor, with fit statistics uniformly worse than those of the one-factor model ($\chi^2/df = 5.35$, TLI = .84, RMSEA = .08). However, a less restricted version of the four-factor model (Figure 2, Model C) that allowed inter-correlations among the four dimension-specific factors (range = .31 to .67) provided much improved fit, with all indicators surpassing the accepted thresholds ($\chi^2/df = 1.73$, TLI = .97, RMSEA = .03). Moreover, since these are nested models, a formal test of the change in chi-square associated with fixing these inter-factor correlations to zero was used to confirm a significant decrement in fit ($\chi^2(6) = 328.47$, $p < .001$).

Second-order factor model—As with the four-factor model, the second-order factor model assumes that covariance among item responses can best be explained by having each indicator load onto one of four dimension-specific factors. However, rather than allowing for correlation between factors, this model posits that the covariance between the four dimension-specific first-order factors can be accounted for by introducing a higher-order general factor (Figure 2, Model D). The second-order factor model provided no discernable improvement in fit over the correlated four-factor model ($\chi^2/df = 1.69$, TLI = .97, RMSEA = .03).

Bifactor model—The bifactor model posits that the covariance among item responses can be best explained by a common factor upon which all indicators load, and dimension-specific factors upon which sub groups of indicators load (Figure 2, Model E). In this way, the bifactor model parses inter-item covariation into that which can be explained by a general syndromal factor representing variance that is shared across all problem behavior indicators, and group-specific factors that model all residual covariation (i.e., not accounted for by the general construct) among behavior-specific subsets of items. For a description comparing the bifactor model to various other factor models see Chen, Hayes, Carver, Laurenceau, and Zhang (2012). Though not nested, fit statistics suggest an improvement in fit over the prior model ($\chi^2/df = 1.58$, TLI = .98, RMSEA = .03). Within this, the best fitting model, all individual item loadings were significant at $p < .001$ (Table 3).

Genetic Association Analyses

While accounting for sex and age, conditional main effects for three candidate genes were assessed by sequentially including individual SNPs as independent variables within the best fitting phenotypic model: the bifactor model. Because this model constrains the covariances between factors to be orthogonal, and indicators load only onto the general factor and one additional dimension-specific factor, this approach yields SNP specific information about prediction of variation in a general factor, as well as variation in unrelated behavior specific factors. The results of regressing the five latent factors on each of the SNPs are summarized in Table 4 in terms of partial R^2 coefficients and p -values for all significant associations.

CHRM2—There were no significant SNP effects of *CHRM2* on *general problem behavior*, but two SNPs each were linearly associated with *alcohol use*, *marijuana use*, and *sexual behavior* factors. Among these associations, only one marker (*rs832990*) approached the adjusted significance threshold for multiple testing, and that was in relation to alcohol use ($R^2 = .016$, $p = .009$). With respect to *disruptive behavior*, five markers were significantly predictive before correction, including three correlated SNPs (*rs11984225*, *rs4407820*, and *rs6970010* with ranges: $R^2 = .014-.027$, $p = .033-.037$).

GABRA2—No SNPs in *GABRA2* were predictive of *general problem behavior*, *alcohol use*, or *marijuana use*. However, three generally distinct SNPs were significantly associated with the *sexual behavior* factor, and four other SNPs with *disruptive behavior*, including one surviving (*rs1442060*: $R^2 = .033$, $p = .002$), and another marginally significant (*rs279851*: $R^2 = .023$, $p = .009$) following the Nyholt-adjusted Bonferroni correction described in the methods section.

OPRM1—There were no genetic associations with *general problem behavior* or *disruptive behavior*, one modest association with *marijuana use*, and one association with *alcohol use* that remained marginal after correcting for multiple testing (*rs613341*: $R^2 = .051$, $p = .005$). In addition, *sexual behavior* was robustly associated with five SNPs representing a single LD block (ranges: $R^2 = .021-.036$, $p = .004-.035$), including one remaining marginal following correction (*rs658156*: $R^2 = .036$, $p = .004$), and another adjacent SNP that survived a Nyholt correction (*rs678122*: $R^2 = .046$, $p = .002$).

Supplementary Genetic Association Analyses

In contrast to the models described above, genotype-based analyses of similar behaviors typically explore associations between one or more genetic variants and a single phenotype of interest; either dimension-specific or broad-based (i.e., incorporating indicators of multiple dimensions of behavior into a single composite or factor score) in nature. In order to determine how genetic influences on more commonly constructed phenotypes for the same set of externalizing behaviors might compare with findings based on the bifactor model, we supplemented our primary analyses by separately regressing five independently derived latent factors (i.e., five separate unidimensional confirmatory factor models, each reflecting a set of indicators from a discrete column in table 3) onto each of the SNPs. Aside from the methods by which the factors were derived, these supplementary analyses were conducted in the same manner as the conditional bifactor analyses. Partial R^2 and p -value

coefficients for significant supplementary associations are presented alongside those of the primary model, in Table 4, and are summarized below.

The relevance of *CHRM2* appeared to differ in the unidimensional models. Most notably, (1) a cluster of correlated SNPs emerged in relation to *alcohol use*, none overlapping with those reflecting a discrete LD block in the primary analyses (ranges: $R^2 = .007-.016$, $p = .018-.05$); (2) two SNPs were associated with *general problem behavior*, one remaining marginally significant (*rs10249472*: $R^2 = .016$, $p = .004$) post correction; and (3) only the most significant marker associated with *disruptive behavior* in the bifactor model remained associated, with a sub-threshold p-value (*rs10249472*: $R^2 = .021$, $p = .003$). *GABRA2* and *OPRM1* were, in contrast, somewhat less associated with unidimensionally derived phenotypic factor scores. Again, three *GABRA2* SNPs were modestly associated with *general problem behavior* before correction, but no other associations emerged. Likewise, in *OPRM1* only one marker was associated with *general problem behavior*, and two with *sexual behavior*, one of which remained marginally significant following the Nyholt correction (i.e., *rs279851*: $R^2 = .012$, $p = .009$).

Discussion

The rapid technological advancements and continued reduction in costs that typify today's era of genomic information have resulted in a proliferation of genetic association studies of both binary and complex behavioral phenotypes. As this body of findings grows at an almost exponential rate, it has become increasingly necessary to characterize the extent to which individual effects (or lack thereof) generalize beyond *a particular operationalization of a construct and the majority populations* typically being studied. The present study attempts to do this with respect to problem behavior, first by establishing the structural nature, and then by examining relationships between the emerging constructs and a set of candidate genes previously shown to influence one or more of the specific problem behavior dimensions, or related externalizing constructs, all within a sample of AA adolescents.

Model fitting comparisons suggest that the bifactor model best accounted for the associations among the item level data in the present sample. Two important implications follow from this finding. First, this indicates the need to recognize and simultaneously assess both general and specific dimensions of problem behavior. Second, unlike how general problem behavior is typically operationalized as the shared variability across a set of *correlated* items or factors reflecting multiple dimensions of behavior (e.g., Figure 2, models A and D, respectively), the bifactor model implies an orthogonal structure among the factors, with each item serving as an indicator of the general syndrome and one specific behavioral dimension. This is consistent with recent evidence characterizing the structural nature of associations among dimensions of externalizing psychopathology (Martel, von Eye, & Nigg, 2010). So, just as it was suggested decades ago, individual behavior problems are linked, though perhaps not in the hierarchical manner proposed by those initially exploring the structural nature of problem behavior via latent variable models (e.g., McGee & Newcomb, 1992).

Given the extant literature on shared genetic liability across domains of problem behavior, it seems natural to question whether these observed phenotypic correlations are due, in part, to the pleiotropic effects of an associated variant (McCarthy & Hirshhorn, 2008), or conversely, whether common genetic influences are due simply to shared phenotypic variance.

Unfortunately, both GWAS and candidate gene studies have generally, if not exclusively, focused on associations between single variants and rather grossly defined stand-alone phenotypes. In the case of GWAS, this was largely because it was thought not to be feasible to increase phenotypic complexity, given logistical computing constraints. However, just as we are now attempting to extend the reach of the candidate gene approach with respect to increasingly complex phenotypes, so too are statistical geneticists actively working toward multivariate phenotypic extensions within the GWAS framework (O'Reilly et al., 2012; van der Sluis, Posthuma, & Dolan, 2013).

It is interesting to note that both *CHRM2* and *GABRA2*, two genes previously implicated via unidimensional models of general externalizing behavior in EA and other diverse samples that statistically control for race (Dick et al., 2008; Dick et al., 2009; Latendresse et al., 2011), showed no evidence of genetic liability with respect to a general factor when fitting a multidimensional bifactor model to data from this sample of AA youths. In fact, the absence of (1) robust associations with the *general problem behavior* factor across correlated SNPs and (2) effects that survive the Nyholt-adjusted Bonferroni correction suggests that these genes may not contribute, at least directly, to the manifestation of a problem behavior syndrome within this sample.

In contrast, all of the candidate genes were associated with the more narrowly defined specific dimensions of problem behavior within the bifactor model. Of particular interest were two variants representing a seemingly novel LD block in *CHRM2*. Despite there being no prior evidence of phenotypic associations with either of these SNPs, one (*rs832990*) remained marginally associated with *alcohol use* in AA youths, suggesting a potentially important region for follow up in subsequent studies.

Likewise, three variants spanning *GABRA2* displayed modest associations with *sexual behavior*. Although none had previously been examined in relation to this dimension of behavior, all have been associated with alcohol phenotypes. However, no associations between *GABRA2* and *alcohol use* were evidenced in this study, perhaps due to lower rates of drinking in this population relative to EA adolescents. Whatever the case, these and surrounding variants should serve as targets in follow up studies.

In addition, two largely independent SNPs within *GABRA2* remained either marginally or significantly associated with *disruptive behavior* after correction, one of which was previously associated with symptoms of conduct disorder in adolescents from the Collaborative Study on the Genetics of Alcoholism sample (Dick et al., 2006). Thus, this particular finding may generalize across adolescent populations, irrespective of ancestral background.

With respect to *OPRM1*, *alcohol use* in this sample was associated with a single SNP deriving from an LD block that includes variants previously associated with alcohol

dependence in EA adults (Zhang et al., 2006), suggesting the possibility that this effect might generalize to AA adolescents (and possibly even AA adults). As such, *OPRM1* should be targeted for further exploration in relation to alcohol phenotypes within AA populations. Even more surprisingly, there was robust evidence of association with *sexual behavior* across the gene, with two uncorrelated variants remaining significant after correction. Previous studies have implicated *OPRM1* as playing a role in social affiliation (Curley, 2011), as well as dispositional and neural sensitivity to social rejection (Way, Taylor, & Eisenberger, 2009), both suggesting a similar mechanism by which this gene might influence sexual behavior. In addition, among those variants associated with *sexual behavior*, one SNP derives from the same LD block as a variant previously associated with heroin dependence among EA (Levrn, Awolesi, Linzy, Adelson, & Kreek, 2011), and another was previously associated with subjective responses to alcohol in adults of American Indian descent (Ehlers et al., 2008).

What then should be made of genetic associations identified via unidimensional models in the literature as compared to those found here? To start, the fact that both significant and robust associations were generally only observed in the primary, but not supplementary analyses at least suggests the possibility that the extant literature on genetic associations with complex behaviors may be complicated by unobserved correlations between and across behavioral dimensions. In addition, a comparison of findings from the primary and supplementary analyses illustrates how significant associations can go undetected when ignoring the complex etiology of these clinically relevant phenotypes.

While the ultimate test resides in replication, there are not yet any comparable studies that have used an extended multidimensional approach. However, a number of issues have been shown to influence reliability in traditional genetic association studies, including phenotypic complexity. In considering numerous multidimensional approaches, statistical geneticists suggest that relative to conventional analyses, latent variable models have the potential to limit multiple comparisons while yielding more precise estimates (Liu, Jorgenson, & Witte, 2005), thereby improving our ability to detect “true” genetic associations by significantly increasing statistical power (Kim, Sohn, & Xing, 2009; van der Sluis, Verhage, Posthuma, & Dolan, 2010).

Notably, developmental scientists have suggested that correlations between various etiological levels of influences (e.g., biological, contextual, social) may be more pronounced in high-risk samples (Cairns & Cairns, 1994; Farmer, Quinn, Hussey, & Holahan, 2001), and that such correlated constraints can propagate misconceptions or stereotypes regarding the risk for developing problem behaviors in racial minority populations. Thus, in as much as genes function to regulate complex behavior, we have to be particularly aware of the potential differences attributable to ancestry. Whereas many studies of genetic association merely control for racial differences, the present sample is among the rare few that explores associations strictly within AA adolescents, a population that is largely understudied with respect to the genetic etiology of behavioral outcomes. Thus, specific associations observed in this study are relatively novel, a fact that would likely remain so even if we had limited our analyses to a more traditional set of unidimensional models. Yet, it warrants mentioning that some degree of ancestral variation has been shown to exist within this sample, and could

serve as a potential source of population stratification. Even so, these multidimensional models should serve as the basis for subsequent replication efforts in AA samples, as well as the impetus for a more thoughtful standard for the modeling of conceptually related phenotypes in genetic association studies irrespective of the demographic makeup.

Among this study's principal limitations is the fact that it was statistically underpowered to reasonably conduct and compare the proposed analyses (i.e., tests of association between three distinct genes and the emerging phenotypes) by sex and/or age, particularly given the large number of SNPs required to adequately characterize variability in genomic regions void of known functional variants in AA populations. So, although prior studies suggest, for example, that development and rates of substance use behaviors (Brady & Randall, 1999), etiology of sexual behavior (Collins, Sutherland, & Kelly-Weeder, 2012), and levels of disruptive behavior problems (Lahey et al., 2000) may vary by sex and/or age, we were only able to statistically control for these factors in our the present analyses. This further highlights the need to identify sufficiently powered independent samples wherein the effects of these genes on complex behaviors can be both compared across distinct populations and replicated within other AA samples, as we strive to expand our understanding of susceptible, yet understudied populations.

Another somewhat related limitation is that of implied invariance within the measurement model that we selected and retained for the analyses. That is, observed indicators may function differentially with respect to age and/or sex, and failure to account for these differences might limit construct validity among the latent externalizing factors, thereby increasing the likelihood of observing spurious genetic associations and/or suppressing meaningful ones. Subsequent research efforts should, therefore, explore potential sources of measurement noninvariance within the bifactor model via the analysis of differential item functioning, and specifically with respect to age and sex.

Notably, clinicians are able to best treat individuals when the etiology of their behavior/disorder is more fully understood. As such, the implications of the present study are two-fold: First, refinement of the phenotypic measurement model increases the precision with which we assess behaviors of interest, and assessment is paramount within the treatment paradigm, especially when individuals present with symptoms reflecting multiple, correlated dimensions. Second, any increased precision in the measurement of behavior further increases the reliability and validity of subsequent associations between genes and the behavior and identifying genes implicated in a given phenotype can serve as an important catalyst with respect to advancing our understanding of the biological processes that underlie it. In turn, researchers will be better prepared to hypothesize about the potential genetic mechanisms by which known risk environments come to influence the phenotype. These are particularly relevant with respect to the biological substrates of problem behavior(s) in AA adolescents, about which relatively little is known.

In conclusion, we provide evidence to support the structural value of modeling both individual and collective dimensions of problem behavior. We further demonstrate the utility of multidimensional models in the identification of genetic variants involved in the etiology of externalizing problems, and illustrate the ramifications of selecting simpler but poorer

fitting models on our capacity to identify and interpret these effects. Phenotypic complexity remains one of the more overlooked obstacles to the replication of genetic associations, and our ability to adequately account for this complexity may serve as a proxy for our success. In addition, we provide initial evidence implicating this set of candidate genes in specific behavioral dimensions that were not detectable at the level of an overarching problem behavior syndrome. While some associations appear to be consistent with findings previously identified in EA or racially diverse samples (e.g., *GABRA2* and *disruptive behavior*; Dick et al., 2006), others are more novel effects that may provide insight into previously unexplored mechanisms: most notably the association between *OPRM1* and *sexual behavior*.

Acknowledgments

This research was supported by the National Institute on Drug Abuse (R01DA025039 to BM) and the National Institute of Alcohol Abuse and Alcoholism (K01AA020333 to SJL and K02AA018755 to DMD).

References

- Achenbach, TM. Manual for the Youth Self-Report and 1991 Profile. Burlington, VT: Department of Psychiatry, University of Vermont; 1991.
- Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*. 2009; 19:1655–1664. <http://dx.doi.org/10.1101/gr.094052.109>. [PubMed: 19648217]
- Allen JP, Leadbeater BJ, Aber JL. The development of problem behavior syndromes in at-risk adolescents. *Development and Psychopathology*. 1994; 6:323–342. <http://dx.doi.org/10.1017/S0954579400004612>.
- Barrett JC, Fry B, Maller J, Daly MJ. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; 21:263–265. <http://dx.doi.org/10.1093/bioinformatics/bth457>. [PubMed: 15297300]
- Bartlett R, Holditch-Davis D, Belyea M. Problem behaviors in adolescents. *Pediatric Nursing*. 2007; 33:13–18. Retrieved from <http://www.pediatricnursing.net>. [PubMed: 17410996]
- Bentler PM, Bonnet DC. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*. 1980; 88:588–606. <http://dx.doi.org/10.1037/0033-2909.88.3.588>.
- Bolland, J. Mobile Youth Survey overview. Birmingham, AL: Department of Health Behavior, School of Public Health, University of Alabama at Birmingham; 2007.
- Brady KT, Randall CL. Gender differences in substance use disorders. *The Psychiatric Clinics of North America*. 1999; 22:241–252. [http://dx.doi.org/10.1016/S0193-953X\(05\)70074-5](http://dx.doi.org/10.1016/S0193-953X(05)70074-5). [PubMed: 10385931]
- Browne, MW., Cudeck, R. Alternative ways of assessing model fit. In: Bollen, KA., Long, JS., editors. *Testing Structural Equation Models*. Beverly Hills, CA: Sage; 1993. p. 136-162. <http://dx.doi.org/10.1177/0049124192021002005>
- Byck, GR., Bolland, J., Dick, DM., Ashbeck, AW., Mustanski, BS. Prevalence of mental health disorders among low-income African American adolescents. *Social Psychiatry and Psychiatric Epidemiology*. 2013. <http://dx.doi.org/10.1007/s00127-013-0657-3>
- Cairns, RB., Cairns, BD. *Lifelines and risks: Pathways of youth in our time*. New York: Harvester Wheatsheaf; 1994.
- Centers for Disease Control and Prevention. Morbidity and Mortality Weekly Report. Jun 8.2012 61:1–168. [Retrieved January 5, 2013] from <http://www.cdc.gov/mmwr/PDF/ss/ss6104.pdf>.
- Chen FF, Hayes A, Carver CS, Laurenceau J-P, Zhang Z. The best of both worlds: Bifactor modeling for simultaneous tests of general and specific facets of multifaceted constructs. *Journal of Personality*. 2012; 80:219–251. <http://dx.doi.org/10.1111/j.1467-6494.2011.00739.x>. [PubMed: 22092195]

- Collins FH, Sutherland MA, Kelly-Weeder S. Gender differences in risky sexual behavior among urban adolescents exposed to violence. *Journal of the American Academy of Nurse Practitioners*. 2012; 24:436–442. <http://dx.doi.org/10.1111/j.1745-7599.2012.00702.x>. [PubMed: 22735067]
- Corley RP, Zeiger JS, Crowley T, Ehringer MA, Hewitt JK, Hopfer CJ, Krauter K. Association of candidate genes with antisocial drug dependence in adolescents. *Drug and Alcohol Dependence*. 2008; 96:90–98. <http://dx.doi.org/10.1016/j.drugalcdep.2008.02.004>. [PubMed: 18384978]
- Curley JP. The mu-opioid receptor and the evolution of mother-infant attachment: Theoretical comment on Higham et al. (2011). *Behavioral Neuroscience*. 2011; 125:273–278. <http://dx.doi.org/10.1037/a0022939>. [PubMed: 21463027]
- Dick DM. Identification of genes influencing a spectrum of externalizing psychopathology. *Current Directions in Psychological Science*. 2007; 16:331–335. <http://dx.doi.org/10.1111/j.1467-8721.2007.00530.x>.
- Dick DM, Aliev F, Wang JC, Gruzza RA, Schuckit M, Kuperman S, Goate A. Using dimensional models of externalizing psychopathology to aid in gene identification. *Archives of General Psychiatry*. 2008; 65:310–318. <http://dx.doi.org/10.1001/archpsyc.65.3.310>. [PubMed: 18316677]
- Dick DM, Bierut L, Hinrichs A, Fox L, Bucholz KK, Kramer J, Foroud T. The role of *GABRA2* in risk for conduct disorder and alcohol and drug dependence across developmental stages. *Behavior Genetics*. 2006; 36:577–590. <http://dx.doi.org/10.1007/s10519-005-9041-8>. [PubMed: 16557364]
- Dick DM, Latendresse SJ, Lansford JE, Budde JP, Goate A, Dodge KA, Bates JE. Role of *GABRA2* in trajectories of externalizing behavior across development and evidence of moderation by parental monitoring. *Archives of General Psychiatry*. 2009; 66:649–657. <http://dx.doi.org/10.1001/archgenpsychiatry.2011.90>. [PubMed: 19487630]
- Dick DM, Latendresse SJ, Riley B. Incorporating genetics into your studies: a guide for social scientists. *Frontier in Psychiatry*. 2011; 2:17. <http://dx.doi.org/10.3389/fpsy.2011.00017>.
- Dick DM, Li T-K, Edenberg HJ, Hesselbrock V, Kramer J, Kuperman S, Foroud T. A genome-wide screen for genes influencing conduct disorder. *Molecular Psychiatry*. 2004; 9:81–86. <http://dx.doi.org/10.1038/sj.mp.4001368>. [PubMed: 14699444]
- Donenberg GR, Emerson E, Bryant FB, Wilson H, Weber-Shifrin E. Understanding AIDS-risk behavior among adolescents in psychiatric care: Links to psychopathology and peer relationships. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2001; 40:642–53. <http://dx.doi.org/10.1097/00004583-200106000-00008>. [PubMed: 11392341]
- Donovan JE, Jessor R. Structure of problem behavior in adolescence and young adulthood. *Journal of Consulting and Clinical Psychology*. 1985; 53:890–904. <http://dx.doi.org/10.1037/0022-006X.53.6.890>. [PubMed: 4086689]
- Donovan JE, Jessor R, Costa FM. Syndrome of problem behavior in adolescence: A replication. *Journal of Consulting and Clinical Psychology*. 1988; 56:762–765. <http://dx.doi.org/10.1037/0022-006X.56.5.762>. [PubMed: 3192793]
- Donovan JE, Jessor R, Costa FM. Adolescent health behavior and conventionality-unconventionality: An extension of problem behavior syndrome. *Health Psychology*. 1991; 10:52–61. [PubMed: 2026131]
- Edenberg HJ, Dick DM, Xuei X, Tian H, Almasy L, Bauer LO, Begleiter H. Variations in *GABRA2*, encoding the $\alpha 2$ subunit of the *GABA_A* receptor, are associated with alcohol dependence and with brain oscillations. *American Journal of Human Genetics*. 2004; 74:705–714. <http://dx.doi.org/10.1086/383283>. [PubMed: 15024690]
- Ehlers CL, Lind PA, Wilhelmsen KC. Association between single nucleotide polymorphisms in the mu opioid receptor gene (*OPRM1*) and self-reported responses to alcohol in American Indians. *BMC Medical Genetics*. 2008; 9:35. <http://dx.doi.org/10.1186/1471-2350-9-35>. [PubMed: 18433502]
- Farmer TW, Quinn MM, Hussey W, Holahan T. The development of disruptive behavioral disorders and correlated constraints: Implications for intervention. *Behavioral Disorders*. 2001; 26:117–130. Retrieved from <http://eric.ed.gov/?id=EJ623150>.
- Hair EC, Park MJ, Ling TJ, Moore KA. Risky behaviors in late adolescence: Co-occurrence, predictors, and consequences. *Journal of Adolescent Health*. 2009; 45:253–261. <http://dx.doi.org/10.1016/j.jadohealth.2009.02.009>. [PubMed: 19699421]

- Hendershot CS, Bryan AD, Feldstein Ewing SW, Claus ED, Hutchison KE. Preliminary evidence for associations of *CHRM2* with substance use and disinhibition in adolescence. *Journal of Abnormal Child Psychology*. 2011; 39:671–681. <http://dx.doi.org/10.1007/s10802-011-9511-9>. [PubMed: 21494862]
- Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis. *Structural Equation Modeling*. 1999; 6:1–55. <http://dx.doi.org/10.1080/10705519909540118>.
- Ittiwut C, Yang BZ, Kranzler HR, Anton RF, Hirunsatit R, Weiss RD, Gelernter J. GABRG1 and GABRA2 variation associated with alcohol dependence in African Americans. *Alcoholism: Clinical and Experimental Research*. 2012; 36:588–593. <http://dx.doi.org/10.1111/j.1530-0277.2011.01637.x>.
- Jessor R. Problem-behavior theory, psychosocial development, and adolescent problem drinking. *British Journal of Addiction*. 1987; 82:331–342. <http://dx.doi.org/10.1111/j.1360-0443.1987.tb01490.x>. [PubMed: 3472582]
- Jessor, R., Donovan, JE., Costa, FM. Beyond adolescence: Problem behavior and young adult development. New York: Cambridge University Press; 1991.
- Jessor, R., Jessor, SL. Problem behavior and psychosocial development: A longitudinal study of youth. New York, NY: Academic Press; 1977.
- Jessor R, Van Den Bos J, Vanderryn J, Costa FM, Turbin MS. Protective factors in adolescent problem behavior: Moderator effects and developmental change. *Developmental Psychology*. 1995; 31:923–933. <http://dx.doi.org/10.1037/0012-1649.31.6.923>.
- Kendler KS, Kalsi G, Holmans PA, Sanders AR, Aggen SH, Dick DM, Gejman PV. Genomewide association analysis of symptoms of alcohol dependence in the molecular genetics of schizophrenia (MGS2) control sample. *Alcoholism: Clinical and Experimental Research*. 2011; 35:963–975. <http://dx.doi.org/10.1111/j.1530-0277.2010.01427.x>.
- Kendler KS, Prescott CA, Myers J, Neale MC. The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry*. 2003; 60:929–937. <http://dx.doi.org/10.1001/archpsyc.60.9.929>. [PubMed: 12963675]
- Kim S, Sohn K-A, Xing EP. A multivariate regression approach to association analysis of a quantitative trait network. *Bioinformatics*. 2009; 25 <http://dx.doi.org/10.1093/bioinformatics/btp218>.
- Lahey BB, Schwab-Stone M, Goodman SH, Waldman ID, Canino G, Rathouz PJ, Jensen PS. Age and gender differences in oppositional behavior and conduct problems: A cross-sectional household study of middle childhood and adolescence. *Journal of Abnormal Psychology*. 2000; 109:488–503. <http://dx.doi.org/10.1037/0021-843X.109.3.488>. [PubMed: 11016118]
- Latendresse SJ, Bates JE, Goodnight JA, Lansford JE, Budde JP, Goate A, Dick DM. Differential susceptibility to adolescent externalizing trajectories: Examining the interplay between *CHRM2* and peer group antisocial behavior. *Child Development*. 2011; 82:1797–1814. <http://dx.doi.org/10.1111/j.1467-8624.2011.01640.x>. [PubMed: 21883161]
- Levrán O, Awolesi O, Linzy S, Adelson M, Kreek MJ. Haplotype block structure of the genomic region of the mu opioid receptor gene (*OPRM1*). *Journal of Human Genetics*. 2011; 56:147–155. <http://dx.doi.org/10.1038/jhg.2010.150>. [PubMed: 21160491]
- Lind PA, Macgregor S, Agrawal A, Montgomery GW, Heath AC, Martin NG, Whitfield JB. The role of *GABRA2* in alcohol dependence, smoking, and illicit drug use in an Australian population sample. *Alcoholism: Clinical and Experimental Research*. 2008; 32:1721–1731. <http://dx.doi.org/10.1111/j.1530-0277.2008.00768.x>.
- Liu X, Jorgenson E, Witte JS. Hierarchical modeling in association studies of multiple phenotypes. *BMC Genetics*. 2005; 6:S104. <http://dx.doi.org/10.1186/1471-2156-6-S1-S104>. [PubMed: 16451560]
- Lonjou C, Zhang W, Collins A, Tapper WJ, Elahi E, Maniatis N, Morton NE. Linkage disequilibrium in human populations. *Proceedings of the National Academy of Sciences of the United States of America*. 2003; 100:6069–6074. <http://dx.doi.org/10.1073/pnas.1031521100>. [PubMed: 12721363]
- Luo X, Kranzler HR, Zhao H, Gelernter J. Haplotypes at the *OPRM1* locus are associated with susceptibility to substance dependence in European-Americans. *American Journal of Medical*

Genetics Part B (Neuropsychiatric Genetics). 2003; 120B:97–108. <http://dx.doi.org/10.1002/ajmg.b.20034>.

- Markon KE, Krueger RF. Categorical and continuous models of liability to externalizing disorders: a direct comparison in NESARC. *Archives of General Psychiatry*. 2005; 62:1352–59. <http://dx.doi.org/10.1001/archpsyc.62.12.1352>. [PubMed: 16330723]
- Martel MM, von Eye A, Nigg JT. Revisiting the latent structure of ADHD: is there a ‘g’ factor? *Journal of Child Psychology and Psychiatry*. 2010; 51:905–914. <http://dx.doi.org/10.1111/j.1469-7610.2010.02232.x>. [PubMed: 20331490]
- McCarthy MI, Hirshhorn JN. Genome-wide association studies: Potential next steps on a genetic journey. *Human Molecular Genetics*. 2008; 17:R156–R165. <http://dx.doi.org/10.1093/hmg/ddn289>. [PubMed: 18852205]
- McGee L, Newcomb MD. General deviance syndrome: Expanded hierarchical evaluations at four ages from early adolescence to adulthood. *Journal of Consulting & Clinical Psychology*. 1992; 60:766–776. <http://dx.doi.org/10.1037/0022-006X.60.5.766>. [PubMed: 1401392]
- McGue M, Iacono WG. The association between early adolescent problem behavior with adult psychopathology. *American Journal of Psychiatry*. 2005; 162:1118–1124. [PubMed: 15930060]
- Monahan KC, Rhew IC, Hawkins JD, Brown EC. Adolescent pathways to co-occurring problem behavior: The effects of peer delinquency and peer substance use. *Journal of Research on Adolescence*. 2013; 24:630–645. <http://dx.doi.org/10.1111/jora.12053>.
- Mustanski, BS., Byck, GR., Bolland, J., Henry, D., Swann, G., Dick, DM. Study protocol: Genes, Environment, and Neighborhood Initiative (GENI): A bioecological study of HIV risk behavior of African American youth living in poverty. Northwestern University, Chicago, Illinois: Department of Medical Social Sciences; 2013. Unpublished Manuscript
- Mustanski B, Starks T, Newcomb ME. Methods for the design and analysis, of relationship and partner effects on sexual health. *Archives of Sexual Behavior*. 2012; 43:21–33. <http://dx.doi.org/10.1007/s10508-013-0215-9>.
- Muthén, LK., Muthén, BO. Mplus user’s guide Sixth Edition. Los Angeles, CA: Authors; 1998–2010.
- Newcomb MD, Abbott RD, Catalano RF, Hawkins JD, Battin-Pearson S, Hill K. Mediation and deviance theories of late high school failure: Process roles of structural strains, academic competence, and general versus specific problem behavior. *Journal of Counseling Psychology*. 2002; 49(2):172–186. <http://dx.doi.org/10.1037/0022-0167.49.2.172>.
- Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *American Journal of Human Genetics*. 2004; 74:765–769. <http://dx.doi.org/10.1086/383251>. [PubMed: 14997420]
- O’Reilly PF, Hoggart CJ, Pomyen Y, Calboli FCF, Elliott P, Jarvelin M-R, Coin LJM. MultiPhen: Joint Model of Multiple Phenotypes Can Increase Discovery in GWAS. *PLoS ONE*. 2012; 7(5):e34861. <http://dx.doi.org/10.1371/journal.pone.0034861>. [PubMed: 22567092]
- Park N, Lee B, Bolland J, Vazsonyi A, Sun F. Early adolescent pathways of antisocial behaviors in poor, inner-city neighborhoods. *Journal of Early Adolescence*. 2008; 28:185–205. <http://dx.doi.org/10.1177/0272431607313591>.
- Reise SP, Morizot J, Hays RD. The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research*. 2007; 16:19–31. <http://dx.doi.org/10.1007/s11136-007-9183-7>. [PubMed: 17479357]
- Resnicow K, Ross-Gaddy D, Vaughan RD. Structure of problem and positive behaviors in African American youths. *Journal of Consulting and Clinical Psychology*. 1995; 63:594–603. <http://dx.doi.org/10.1037/0022-006X.63.4.594>. [PubMed: 7673537]
- Risch N, Burchard E, Ziv E, Tang H. Categorization of humans in biomedical research: genes, race and disease. *Genome Biology*. 2002; 3:1–12. <http://dx.doi.org/10.1186/gb-2002-3-7-comment2007>.
- Seglem KB, Torgersen S, Ask H, Waaktaar T. Weak etiologic links between control and the externalizing behaviors delinquency and substance use in adolescence. *Personality and Individual Differences*. 2015; 75:179–184. <http://dx.doi.org/10.1016/j.paid.2014.11.036>.
- Sartor CE, Grant JD, Bucholz KK, Madden PA, Heath AC, Agrawal A, Lynskey MT. Common genetic contributions to alcohol and cannabis use and dependence symptomatology. *Alcoholism: Clinical*

and Experimental Research. 2010; 34:545–554. <http://dx.doi.org/10.1111/j.1530-0277.2009.01120.x>.

- Shaffer D, Fisher P, Lucas C, Dulcan MK, Schwab-Stone M. NIMH Diagnostic Interview Schedule for Children, Version IV (NIMH DISC-IV): Description, differences from previous versions and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2008; 39:28–38. <http://dx.doi.org/10.1097/00004583-200001000-00014>.
- Smelson D, Yu L, Buyske S, Gonzalez G, Tischfield J, Deutsch CK, Ziedonis D. Genetic association of *GABA-A* alpha-2 and mu opioid receptor with cocaine cue-reactivity: Evidence for inhibitory synaptic neurotransmission involvement in cocaine dependence. *American Journal of Addictions*. 2012; 21:411–415. <http://dx.doi.org/10.1111/j.1521-0391.2012.00253.x>.
- Stallings MC, Corley R, Dennehey B, Hewitt JK, Krauter KS, Lessem JM, Crowley TJ. A genome-wide search for quantitative trait loci that influence antisocial drug dependence in adolescence. *Archives of General Psychiatry*. 2005; 62:1042–1051. <http://dx.doi.org/10.1001/archpsyc.62.9.1042>. [PubMed: 16143736]
- Tabachnick, BG., Fidell, LS. Using multivariate statistics. Fifth Edition. New York: Allyn & Bacon; 2007.
- Tackett JL. Toward and externalizing spectrum in DSM-V: Incorporating developmental concerns. *Child Development Perspectives*. 2010; 4:161–167.
- Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nature Reviews Genetics*. 2002; 3:391–397. <http://dx.doi.org/10.1038/nrg796>.
- Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*. 1973; 38:1–10. <http://dx.doi.org/10.1007/BF02291170>.
- United States Census Bureau. [Retrieved December 10, 2004] 2002 Census. 2002. from <http://www.census.gov>
- van den Oord EJCG, Jiang Y, Riley BP, Kendler KS, Chen X. FP-TDI SNP scoring by manual and statistical procedures: A study of error rates and types. *Biotechniques*. 2003; 34:610–624. [PubMed: 12661166]
- van der Sluis S, Posthuma D, Dolan CV. TATES: Efficient Multivariate Genotype-Phenotype Analysis for Genome-Wide Association Studies. *PLoS Genetics*. 2013; 9(1):e1003235. <http://dx.doi.org/10.1371/journal.pgen.1003235>. [PubMed: 23359524]
- van der Sluis S, Verhage M, Posthuma D, Dolan CV. Phenotypic complexity, measurement bias, and poor phenotypic resolution contribute to the missing heritability problem in genetic association studies. *PLoS ONE*. 2010; 5:e13929. <http://dx.doi.org/10.1371/journal.pone.0013929>. [PubMed: 21085666]
- Vazsonyi AT, Chen P, Jenkins DD, Burcu E, Torrente G, Sheu C-J. Jessor's problem behavior theory: Cross-national evidence from Hungary, the Netherlands, Slovenia, Spain, Switzerland, Taiwan, Turkey, and the United States. *Developmental Psychology*. 2010; 46:1779–1791. <http://dx.doi.org/10.1037/a0020682>. [PubMed: 20873922]
- Verweij KJH, Zietsch BP, Bailey JM, Martin NG. Shared aetiology of risky sexual behaviour and adolescent misconduct: genetic and environmental influences. *Genes, Brain, and Behavior*. 2009; 8:107–113. <http://dx.doi.org/10.1111/j.1601-183X.2008.00456.x>.
- Way BM, Taylor SE, Eisenberger NI. Variation in the μ -opioid receptor gene (*OPRM1*) is associated with dispositional and neural sensitivity to social rejection. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106:15079–15084. <http://dx.doi.org/10.1073/pnas.0812612106>. [PubMed: 19706472]
- Wheaton B, Muthén B, Alwin DF, Summers G. Assessing reliability and stability in panel models. *Sociological Methodology*. 1977; 8:84–136. <http://dx.doi.org/10.2307/270754>.
- Zhang H, Luo X, Kranzler HR, Lappalainen J, Yang B-Z, Krupitsky E, Gelernter J. Association between two μ -opioid receptor gene (*OPRM1*) haplotype blocks and drug or alcohol dependence. *Human Molecular Genetics*. 2006; 15:807–819. <http://dx.doi.org/10.1093/hmg/ddl024>. [PubMed: 16476706]

Figure 1.

LD structure within the GENI sample for (a) *CHRM2*, (b) *GABRA2*, and (c) *OPRM1*, with individual coefficients reflecting pairwise R^2 values.

Model A

Model B

Model C

Model D

Model E

Figure 2.

Confirmatory factor models used to assess the dimensionality of problem behavior: (Model A) One-factor model, (Model B) Four-factor model, (Model C) Correlated Four-factor model, (Model D) Second-order factor model, (Model E) Bifactor model.

Table 1

Targeted Single Nucleotide Polymorphisms (SNPs)

<i>CHRM2</i>			<i>GABRA2</i>			<i>OPRM1</i>		
SNP ^a	Location ^b	MAF ^c	SNP ^a	Location ^b	MAF ^c	SNP ^a	Location ^b	MAF ^c
<i>rs10249472</i>	136049718	.26	<i>rs497068</i>	45945434	.35	<i>rs12205732</i>	154400626	.10
<i>rs6977141</i>	136091394	.35	<i>rs548583</i>	45958101	.32	<i>rs1799971</i>	154402490	.02
<i>rs10954546</i>	136102622	.42	<i>rs526805</i>	45958148	.36	<i>rs553202</i>	154406510	.44
<i>rs11984255</i>	136141629	.42	<i>rs537134</i>	45966025	.25	<i>rs524731</i>	154416785	.12
<i>rs4407820</i>	136161629	.42	<i>rs279871</i>	46000490	.23	<i>rs3778150</i>	154425351	.19
<i>rs6970010</i>	136166257	.37	<i>rs1808851</i>	46006204	.28	<i>rs10457090</i>	154432766	.07
<i>rs123333547</i>	136182323	.25	<i>rs279858</i>	46009350	.22	<i>rs9478503</i>	154434368	.21
<i>rs6967120</i>	136186671	.30	<i>rs279851</i>	46014748	.44	<i>rs589046</i>	154434831	.44
<i>rs10246433</i>	136191614	.38	<i>rs1561774</i>	46015347	.35	<i>rs3778157</i>	154447394	.19
<i>rs978437</i>	136264718	.23	<i>rs11941602</i>	46018451	.21	<i>rs10485057</i>	154454948	.18
<i>rs7782965</i>	136274673	.23	<i>rs279843</i>	46019961	.43	<i>rs562859</i>	154456266	.40
<i>rs7800170</i>	136274860	.34	<i>rs279845</i>	46024480	.26	<i>rs511420</i>	154465725	.21
<i>rs1455858</i>	136282243	.19	<i>rs183960</i>	46025645	.33	<i>rs9322447</i>	154466013	.35
<i>rs1378646</i>	136285541	.23	<i>rs1440130</i>	46028010	.29	<i>rs681243</i>	154469433	.48
<i>rs1824024</i>	136294234	.32	<i>rs279826</i>	46028966	.48	<i>rs504932</i>	154472161	.18
<i>rs3244582</i>	136301147	.26	<i>rs279828</i>	46029567	.27	<i>rs512053</i>	154481209	.01
<i>rs2061174</i>	136311940	.48	<i>rs279836</i>	46033827	.23	<i>rs658156</i>	154483218	.42
<i>rs324650</i>	136344201	.32	<i>rs1442060</i>	46060824	.41	<i>rs645027</i>	154483822	.09
<i>rs7779921</i>	136365936	.21	<i>rs1442062</i>	46071833	.30	<i>rs644261</i>	154483943	.14
<i>rs7796601</i>	136444656	.22	<i>rs11503014</i>	46085622	.25	<i>rs613341</i>	154484971	.09
<i>rs832990</i>	136491912	.32				<i>rs616585</i>	154485574	.27
<i>rs11971038</i>	136492737	.31				<i>rs10485058</i>	154486907	.06
<i>rs832988</i>	136492957	.38				<i>rs497315</i>	154489237	.24
<i>rs833018</i>	136522917	.26				<i>rs678122</i>	154491795	.42
<i>rs833016</i>	136525108	.36						

Note

^a markers shown as rs numbers from the dbSNP database;

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

q locations are derived from estimates of chromosomal position in the dbSNP database (build 36.3) and expressed in base pairs;
 c sample-specific minor allele frequencies.

Table 2

Fit Statistics for Unconditional Confirmatory Factor Models

Model	χ^2	df	χ^2/df	TLI	RMSEA (90% CI)	Parameters Estimated
A. One-factor model	2395.51	464	5.16	.85	.081 (.078, .085)	106
B. Four-factor model	2483.11	464	5.35	.84	.083 (.080, .086)	106
C. Correlated four-factor model	793.64	458	1.73	.97	.034 (.030, .038)	112
D. Second-order factor model	778.38	460	1.69	.97	.033 (.029, .037)	110
E. Bifactor model	684.14	432	1.58	.98	.030 (.026, .035)	138

Note: For all chi-squares statistics, N = 590; TLI = Tucker-Lewis index; RMSEA = root-mean-square error of approximation; BIC = Bayesian information criteria.

Table 3

Standardized Loadings for Indicators in the Bifactor Model

	Factors				
	G	AU	MU	RS	DB
Alcohol Use: In the LAST 12 MONTHS...					
have you had six or more drinks?	.78	.51			
how many days did you drink alcohol?	.66	.67			
how many drinks did you usually have each time?	.64	.71			
how many days did you drink 5 or more drinks in a row?	.66	.62			
Marijuana Use: In the LAST 12 MONTHS...					
have you used marijuana in the last year?	.89		.45		
have you used marijuana six or more times?	.86		.47		
how many times did you use marijuana?	.86		.41		
Sexual Behavior: In your LIFETIME...					
have you ever had oral sex	.50			.35	
have you ever had vaginal sex	.47			.88	
have you ever had anal sex	.25 ^a			.41	
how many people have you had vaginal or anal sex with?	.45			.76	
how many partners did you NOT use a condom or rubber with during vaginal or anal sex?	.37			.57	
Disruptive Behavior: Now or within the last 6 MONTHS...					
i argue a lot	.21				.71
i disobey my parents	.40				.63
i disobey at school	.38				.73
i am stubborn	.21				.40
i have a hot temper	.22				.52
i am mean to others	.20 ^b				.72
i destroy things belonging to others	.47				.59
i don't feel guilty after doing something i shouldn't	.26				.41
i break rules at home, school, or elsewhere	.39				.72
i get in many fights	.32				.54
i hang around kids who get in trouble	.37				.52

	Factors				
	G	AU	MU	RS	DB
i lie or cheat	.42				.55
i physically attack people	.35				.68
i run away from home	.32				.41
i set fires	.53				.46
i steal at home	.36				.49
i steal from places other than home	.57				.37
i swear or use dirty words	.50				.50
i threaten to hurt people	.54				.61
i cut classes or skip school	.59				.30

Note. G = general problem behavior; AU = alcohol use; MU = marijuana use; SB = sexual behavior; DB = disruptive behavior;

^a $p = .007$;

^b $p = .004$; all other loadings were significant at $p = .001$.

Table 4

Significant Main Effects of Markers, by Gene and Analysis

Models:	Primary Analyses					Supplementary Analyses				
	Bifactor Model					Unidimensional Confirmatory Factor Models				
	G	AU	MU	SB	DB	G	AU	MU	SB	DB
rs10249472					.014(.017)	.016(.004)				.021(.003)
rs6977141										
rs10954546										
rs11984255					.014(.036)					
rs4407820					.026(.037)					
rs6970010					.027(.033)					
rs12333547										
rs6967120										
rs10246433										
rs978437						.007(.050)				
rs7782965						.014(.029)				
rs7800170						.006(.033)	.016(.018)			
rs1455858						.015(.038)				
rs1378646						.013(.047)				
rs1824024										
rs324582										
rs2061174										

Models:	Primary Analyses						Supplementary Analyses					
	Bifactor Model						Unidimensional Confirmatory Factor Models					
	G	AU	MU	SB	DB	G	AU	MU	SB	DB		
CHRM2												
rs324650				.021(.027)	.021(.044)							
rs7779921												
rs7796601								.023(.030)				
rs832990		.016(.009)	.002(.016)									
rs11971038										.011(.024)		
rs832988			.022(.026)	.024(.028)								
rs833018		.038(.035)										
rs833016												
GABRA2	G	AU	MU	SB	DB	G	AU	MU	SB	DB		
rs497068												
rs548583												
rs526805												
rs537134									.011(.028)			
rs279871									.013(.035)			
rs1808851												
rs279858				.027(.036)						.017(.025)		
rs279851					.023(.009)							
rs1561774					.014(.049)							
rs11941602												

Models:	Primary Analyses						Supplementary Analyses					
	Bifactor Model						Unidimensional Confirmatory Factor Models					
	G	AU	MU	SB	DB		G	AU	MU	SB	DB	
<i>CHRM2</i>												
<i>rs279843</i>				.026(.026)								
<i>rs279845</i>												
<i>rs183960</i>					.010(.024)							
<i>rs1440130</i>												
<i>rs279826</i>												
<i>rs279828</i>												
<i>rs279836</i>												
<i>rs1442060</i>					.033(.002)							
<i>rs1442062</i>				.022(.041)								
<i>rs11503014</i>												
<i>OPRMI</i>	G	AU	MU	SB	DB		G	AU	MU	SB	DB	
<i>rs12205732</i>												
<i>rs1799971</i>												
<i>rs553202</i>												
<i>rs524731</i>												
<i>rs3778150</i>												
<i>rs10457090</i>												
<i>rs9478503</i>												
<i>rs589046</i>												

Models:	Primary Analyses						Supplementary Analyses					
	Bifactor Model						Unidimensional Confirmatory Factor Models					
	G	AU	MU	SB	DB		G	AU	MU	SB	DB	
<i>CHRM2</i>												
<i>rs3778157</i>											.012(.009)	
<i>rs10485057</i>												
<i>rs562859</i>												
<i>rs511420</i>									.009(.047)			
<i>rs9322447</i>												
<i>rs681243</i>						.033(.010)						
<i>rs504932</i>						.028(.026)						
<i>rs512053</i>												
<i>rs658156</i>						.036(.004)						
<i>rs645027</i>												
<i>rs644261</i>												
<i>rs613341</i>		.051(.005)				.021(.035)						
<i>rs616585</i>			.031(.050)			.034(.030)						
<i>rs10485058</i>												
<i>rs497315</i>												
<i>rs678122</i>				.046(.002)							.008(.020)	

Note. G = general problem behavior; AU = alcohol use; MU = marijuana use; SB = sexual behavior; DB = disruptive behavior. Coefficients reflect R^2 change attributed to a genetic marker and associated p -values (in parentheses). **Bold face type** denotes effects maintaining significance after accounting for the estimated number of independent marker loci. Markers are shown as rs numbers from the dbSNP database and ordered by chromosomal position as denoted in build 36.3.