Integrating Longitudinal Twin Research and Measured Genetic Variation to Test Developmental Theories of Behavioral Disinhibition and Adolescent Substance Use

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Scott Ian Vrieze

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

William Merrill Grove, Matt McGue

August 2012

Acknowledgements

Dedication

To Danielle and Maria.

Table of Contents

List of Tables

List of Figures

Introduction

This dissertation is composed of three studies that evaluate the genetic etiology of substance use, and test the notion that behavioral disinhibition contributes to the initiation and maintenance of substance use in adolescents and young adults. Here we briefly describe each study.

1. Study 1 is an attempt to understand longitudinal change in the relationships among substance use disorders. We find that substance use symptom counts become less and less related to one another as individuals age from 14 to 29, suggesting that early adolescent use is driven more by general vulnerability to experiment and use, whereas young adults tend to specialize in their substance use habits (e.g., by restricting use to fewer substances). More specifically, we evaluated the pattern of correlations among nicotine, alcohol, and marijuana abuse and dependence symptom counts as well as their underlying genetic and environmental influences in a community-representative twin sample (N=3762). Symptoms were assessed at ages 11, 14, 17, 20, 24, and 29. A single common factor was used to model the correlations among symptom counts at each age. Age-related changes in the influence of this general factor were examined by testing for differences in the mean factor loading across time. Mean levels of abuse/dependence symptoms increased throughout adolescence, peaked around age 20, and declined from age 24 to 29. The influence of the general factor was highest at ages 14 and 17, but decreased from age 17 to 24. Genetic influences of the general factor declined considerably with age, along with an increase in non-

shared environmental influences. Adolescent substance abuse/dependence is largely a function of shared etiology. As individuals age, symptoms are increasingly influenced by substance-specific etiological factors. Heritability analyses showed that the generalized risk is primarily influenced by genetic factors in adolescence, but non-shared environmental influences increase in importance as substance dependence becomes more specialized in adulthood.

2. Study 2 is an investigation of substance use heritability and the impact of single nucleotide polymorphisms. We used a large community-representative sample (N = 7188) to investigate the genetic and environmental relationships among measures of Behavioral Disinhibition, Nicotine Use, Alcohol Consumption, Alcohol Dependence, and Drug Use. First, using a subsample of twins (N = 2877), we used standard twin models to estimate the additive genetic, shared environmental, and non-shared environmental contributions to these five traits. Heritabilities ranged from .42-.58. Shared environmental effects ranged from .12-.24. Phenotypic correlations among the five traits were largely attributable to shared genetic effects. Second, we used Genome-wide Complex Trait Analysis (GCTA) to estimate as a random effect the aggregate genetic effect attributable to 527,829 common SNPs. The aggregated SNPs explained 10-30% of the variance in the traits. Third, a genome-wide scoring approach summed the actual SNPs, creating a SNP-based genetic risk score for each individual. After 10-fold internal cross-validation, the SNP sumscore correlated with the traits at .025 to .060, indicating small but detectable effects. SNP sumscores generated on one

trait correlated at approximately the same magnitude with other traits, indicating

detectable pleiotropic effects among these traits. Behavioral disinhibition thus

shares genetic etiology with measures of substance use, and this relationship is

detectable at the level of measured genomic variation.

3. Related to the previous study, we attempted to use a SNP score to evaluate gene

by development interaction in cigarettes smoked per day (CPD) in a longitudinal

community-representative sample (N=3231) of Caucasian twins measured at ages

14, 17, 20, and 24. Biometric heritability analyses show strong heritabilities and

shared environmental influences, as well as cross-age genetic and shared-

environmental correlations. SNPs previously associated with CPD according to

meta-analysis were summed to create a SNP score. At best, the SNP score

accounted for 1% of the variance in CPD. The results suggest developmental

moderation with a larger significant SNP score effect on CPD at age 20 and 24,

and smaller non-significant effect at age 14 and 17. These results are consistent

with the notion that nicotine-specific genetic substance use risk is less important

at younger ages, and becomes more important as individuals age into adulthood.

In a complementary analysis, the same nicotine-relevant SNP score was unrelated

to the frequency of alcohol use at ages 14, 17, 20, or 24. These results indicate

that the SNP score is specific to nicotine in this small sample and that increased

exposure to nicotine at ages 20 and 24 does not influence the extent of concurrent

or later alcohol use. Increased sample sizes and replication or meta-analysis are

necessary to confirm these results. The methods and results illustrate the

importance and difficulty of considering developmental processes in

understanding the interplay of genes and environment.

Study 1. Genetic Influence on the Co-Occurrence of Alcohol, Marijuana, and Nicotine Dependence Symptoms Declines from Age 14 to 29.

Epidemiological studies have documented high rates of comorbidity among alcohol, nicotine, and illicit drug dependence disorders (Kendler, Jacobson, Prescott, & Neale, 2003; Kendler, Prescott, Myers, & Neale, 2003; Kessler, Chiu, Demler, Merikangas, & Walters, 2005; Krueger et al., 2002; Moffitt et al., 2010). Comorbidity suggests that these nosologically distinct disorders are caused, in part, by common etiological processes (Krueger & Markon, 2006; Meehl, 1992, 2001). In substance use disorders, the etiology shared among disorders indexed by comorbidity has been referred to as "externalizing" or "disinhibitory psychopathology" (Iacono, Malone, & McGue, 2008).

While comorbidity among substance use disorders has been extensively studied cross-sectionally—typically for lifetime prevalence rates—few studies have examined how comorbidity changes over time, an important topic given the large changes in incidence over the lifespan. Specifically, substance use tends to emerge in middle adolescence, increases substantially throughout adolescence, peaks in the early 20's, plateaus and then decreases in the late 20's (Young et al., 2002). Whether the rates of comorbidity among substance use disorders remain consistently high, however, is unknown. For example, rates of comorbidity might decline suggesting individuals begin to specialize in their substance use over time.

To investigate this, we used a large, longitudinal twin study to examine change in correlations among substance use disorder symptom counts over time. Assessments coincided with key developmental transitions in substance use including prior to initiation (age 11), initiation (age 14), regular use (age 17), heavy use and dependence (age 20 and 24), and the period when individuals decrease their use or exhibit patterns of persistent substance use problems (age 29). First, we examined patterns of mean-level change in nicotine, alcohol, and marijuana dependence symptoms. Second, we fit a single common factor model to account for the correlations among incident symptom counts of nicotine, alcohol, and marijuana at each age. This common factor model then allowed us to test for changes in the contribution of common and specific etiological influences on these disorder symptom counts across time. Third, we used standard twin models to estimate genetic and environmental influences on the common factor over time.

In addition to using the full sample, we also fit separate models to a subsample of early-onset users. We did so because estimates of correlations at earlier ages may be due to a minority of high-risk individuals who tend to exhibit dependence symptoms for multiple substances. Within this high-risk sample then correlations among dependence symptoms may remain high into adulthood. This could be obscured, however, in analyses of the full sample as substance use becomes more normative in adulthood. That is, if the correlations among substance use disorders are lower for the larger group of later-onset users, what is actually an artifact of early- versus later-onset use would appear to be an overall decline in the correlations among substance use disorders. We thus performed

our analyses both for the full sample, and in a subsample of participants who had at least one symptom by their age-17 assessment, prior to the age at which any of these substances become legal in the United States.

Method
Sample

Participants (N=3762; 52% female) were drawn from the Minnesota Twin Family Study, a community-representative longitudinal study of Minnesota families (Iacono & McGue, 2002). The younger twin cohort (N=2510; 51% female) was first assessed at age 11 during the years 1988-2005. The older cohort was first assessed at age 17 (N=1252; 54% female) during the years 1989-1996. Members of the 11-year old cohort were invited to participate in follow-up assessments at age 14 and 17, and all twins were invited to participate in follow-up assessments at age 20, 24, and 29. Cohorts were combined for all analyses. Participants received modest payments for their assessments. Written assent or consent was obtained from all participants, including the parents of minor children, and all study protocols were approved by the University of Minnesota IRB.

Additional analyses were conducted with a subsample of participants who had at least one symptom by their age-17 assessment for nicotine, alcohol, or marijuana dependence. This resulted in an "early-use" subsample of 580 males and 486 females.

Pooling across cohorts in the full sample, the actual ages of assessment were 11.8 (SD=0.4), 14.9 (0.6), 17.8 (0.7), 21.1 (0.8), 25.0 (0.9), and 29.5 (0.7) years. Participation rates ranged from 87.3% to 93.6% for the follow-up assessments. To examine attrition,

we compared 17-year-olds who did versus did not complete the adult assessments at age

20, 24, and 29. For males, Cohen's d for mean differences in age 17 dependence

symptoms between those who did (N=1570) versus did not (N=238) complete the later

assessments were .00, -.08, and .09 for nicotine, alcohol, and marijuana, respectively.

For females, Cohen d's for similar comparisons were -.19, -.01, and .13 (all p's>.05).

Measures

Diagnostic symptom counts were obtained during in-person interviews with

trained interviewers. In a consensus process, graduate students and staff with advanced

training in clinical assessment reviewed cases to verify symptom presence.

At the age-11 and age-14 assessments, participants were assessed for DSM-III-R

(American Psychiatric Association, 1987) nicotine dependence, alcohol

dependence/abuse, and marijuana dependence/abuse using the Diagnostic Interview for

Children and Adolescents (*DICA-R*; Welner, Reich, Herjanic, Jung, & Amado, 1987).

All later assessments used a modified version of the Substance Abuse Module (SAM;

Robins, Babor, & Cottler, 1987) of the Composite International Diagnostic Interview

(CIDI; Robins et al., 1988) to assess DSM-III-R symptoms of substance use disorders.

Abuse and dependence symptoms were collapsed for alcohol and marijuana. Mother

reports of their children's symptoms were also obtained at ages 11, 14, and 17. The

follow-up assessments at each age covered the interval elapsing since the last assessment.

A "best-estimate" approach (Leckman, Scholomskas, Thompson, Belanger, & Weisman,

1982) was used whereby a symptom was considered present if reported by either the

child or mother. Diagnostic inter-rater reliability of substance use disorders was greater

than .91 (Iacono, Carlson, Taylor, Elkins, & McGue, 1999). To rule out possible

informant effects, all analyses were repeated using only the child as the informant. The

pattern of results was identical.

Analysis of Change in Comorbidity

Bivariate correlations were computed using Spearman's rho statistic, a rank-order

statistic robust to departures from bivariate normality (Caruso & Cliff, 1997).

Confirmatory factor analysis (Brown, 2006) was used to model the pattern of correlations

among substance use disorders over time. For each assessment age, a single factor was

fit to account for correlations among symptom counts of alcohol, nicotine, and marijuana

dependence. Due to prohibitively low variance, symptoms at age 11 were not included in

the model. This resulted in a model with five general factors, one for each age of

assessment, and each representing the covariance among substance use disorders at each

age.   All factors were allowed to correlate, and all same-drug residuals were allowed to

correlate across ages to account for the within-person correlated nature of the longitudinal

data.   Loadings were standardized. As such, the variance of the symptom count

variables can be modeled as a function of the general and residual (substance specific)

factors:

$$\text{Var(Symptom Count)} = (\text{Factor Loading})^2 \times \text{Var(Symptom Count)} + (\text{Residual})^2 , \quad \textbf{(1)}$$

where "Var" denotes variance. Since the variance of the general factor was set to 1, the

variance in a standardized symptom count accounted for by the general factor is simply

the square of the factor loading. To obtain a single estimate of comorbidity for each age

of assessment (i.e., a single estimate of the variance in the three symptom counts

9

accounted for by the general factor), we calculated the mean squared factor loadings at each age. The mean squared loading provides a reasonable metric to test for changes in correlations, because it is directly proportional to the magnitude of correlations among the symptom count variables (i.e., higher correlations among the symptom counts results in greater mean variance in the symptom counts accounted for by the general factor). To test for significant differences, we constrained the mean squared loading to be the same across assessment waves, and conducted likelihood ratio tests to evaluate the change in model fit according to standard practice (Miller & Neale, 1995).

In addition, because the sample was composed of twins, we used standard twin modeling to decompose the general factor variances into three components: additive genetic (A), shared environment (C), and non-shared environment (E). All analyses were conducted in the same model rather than separate phenotypic and biometric models. Because we set the variance of the general factor to 1, the sum of the A, C, and E variance components of the general factors is also 1. To obtain the average variance in the symptom counts accounted for by genetic influences on the factor, one simply substitutes the term "Var(General Factor)" in the above equation with the corresponding genetic, shared environmental, or non-shared environmental component of the general factor variance. For example, if the age 17 mean squared loading was .4, then that would indicate that 40% of the variance in the symptom counts was due to the general factor. If the additive genetic component A of the age-17 factor was .7, then the genetic influence of the general factor onto the symptom counts would be $.4 \times .7 = .28$. That is, 28% of the symptom count variance would be due to the genetic influence of the factor.

Note that investigating the biometric decomposition of symptom count

correlations is only applicable to the full sample, because the subsample analysis of

individuals symptomatic by age 17 was, by definition, a within-individual analysis and

disregards unaffected or later-affected (after age-17) co-twins. Resulting biometric

decompositions would be difficult to interpret because included in the cross-twin

correlations are only those pairs of twins both of whom were symptomatic by age 17.

Analyses were conducted using *R* version 2.10.1 (R Development Core Team,

2011). OpenMx version 1.0.6 (Boker et al., 2011) was used for the factor analysis.

Missing data was handled using full information maximum likelihood. Model fit was

evaluated using χ^2 tests and the difference in the Akaike Information Criterion (Vrieze,

2012) between the saturated and alternative models (positive values indicate the

alternative provides better fit), separately for the four samples under study. We also

report the root mean squared error of approximation, where values of .06 or less indicate

a very good fit (Hu & Bentler, 1999). The χ^2 test provides an exact test of model fit, but

is sensitive to sample size and the magnitude of correlations among measures, and is

always significant (p<.05) when statistical power is high (e.g., in large samples). It is

often significant even when the model provides an accurate and useful representation of

the data (Bentler, 1990). As such, other fit indices have been developed and are primarily

used to evaluate model fit. The root mean square error of approximation attempts to

correct deficiencies of the χ^2 by adjusting for the degrees of freedom and sample size.

Conventional RMSEA cutoffs are .08 for good fit and .05 for very good fit. The Akaike

Information Criterion has strong theoretical properties in that the selected model is expected to fit best upon cross-validation (Vrieze, 2012).

Results

Longitudinal trends in symptom counts are somewhat different for males and females (Table 1 and Figure 1). For males, symptom counts significantly increase from age 11 to late adolescence, peak during the early 20's, and decline thereafter. For females there is a more prolonged plateau during the late teens and early 20's with a marked drop only by age 29. Mean symptom counts increased more rapidly for males than females, with males maintaining higher mean-level symptoms after age 14. Variances also increased during adolescence, and decreased during the late 20's.

Table 1. Nicotine, Alcohol, and Marijuana Use and Misuse from Age 11 to 29.

Target Age of Assessment	Substance	N		% Ever Used Substance		Prevalence Of Dependence(%)		Prevalence of Subclinical Dependence (% with 1 or 2 Symptoms)		Mean Symptom Count		SD of Symptom Count		Sex Differences in Mean Symptom Count[a]	Sex Differences in Variance of Symptom Count[b]
		M	F	M	F	M	F	M	F	M	F	M	F	Effect Size	M/F Ratio
11	Nicotine	1233	1282	9	4	0.2	0.1	0.3	0.0	0.02	0.00	0.25	.08	N/A	N/A
	Alcohol	1233	1281	8	2	0.0	0.0	0.1	0.0	0.00	0.00	0.03	0.00	N/A	N/A
	Marijuana	1233	1282	1	0.4	0.0	0.0	0.0	0.0	0.00	0.00	0.03	0.00	N/A	N/A
14	Nicotine	1148	1191	35	27	4.5	4.6	3.5	2.8	0.22	0.31	0.99	1.15	-0.08	0.74***
	Alcohol	1132	1191	53	41	1.7	0.6	3.1	3.2	0.08	0.08	0.54	0.55	0.00	0.96
	Marijuana	1132	1191	13	10	1.7	1.4	2.4	1.5	0.09	0.08	0.67	0.62	0.02	1.17*
17	Nicotine	1129	1644	69	54	17.1	12.5	11.2	8.6	0.96	0.72	1.82	1.62	0.14**	1.26***
	Alcohol	1420	1644	79	77	11.1	4.9	23.6	13.3	0.74	0.37	1.55	1.19	0.27***	1.70***
	Marijuana	1423	1644	34	30	9.2	3.2	9.5	5.7	0.59	0.26	1.67	1.02	0.24***	2.68***
20	Nicotine	1393	1331	87	72	27.4	17.0	15.1	12.0	1.50	0.90	1.95	1.63	0.33***	1.43***
	Alcohol	1101	1331	96	96	24.6	6.2	30.0	15.6	1.56	0.44	1.93	1.12	0.73***	2.97***

13

	Marijuana	1111	1330	55	45	11.8	3.8	13.4	6.8	0.84	0.27	1.69	0.97	0.42***	3.04***
	Nicotine	1009	1213	90	77	26.4	17.3	14.1	12.5	1.34	0.90	1.83	1.60	0.26***	1.31***
24	Alcohol	1148	1213	98	98	23.0	6.9	32.0	17.8	1.39	0.52	1.76	1.23	0.58***	2.05***
	Marijuana	1143	1211	60	50	7.4	3.1	13.2	5.4	0.56	0.20	1.39	0.85	0.31***	2.67***
	Nicotine	957	636	90	79	23.1	15.7	16.2	10.9	1.22	0.79	1.77	1.55	0.26***	1.30**
29	Alcohol	957	636	98	97	12.9	4.5	27.4	11.2	0.89	0.31	1.52	0.95	0.44***	2.56***
	Marijuana	955	635	60	46	5.5	0.6	8.0	3.7	0.42	0.09	1.25	0.52	0.32***	5.78***

The prevalence rates for each disorder and the proportion of individuals who had ever used each substance for each age are reported in Table 1. The mean participant age of initiation was 14.4 (SD=3.4) for nicotine, 15.5 (SD=2.6) for alcohol use without parental permission, and 16.5 (SD=2.5) years for marijuana. The mean age of symptom onset was 17.5 (SD = 2.7) for nicotine, 18.2 (SD=4.4) for alcohol, and 17.2 (SD = 2.3) for marijuana.

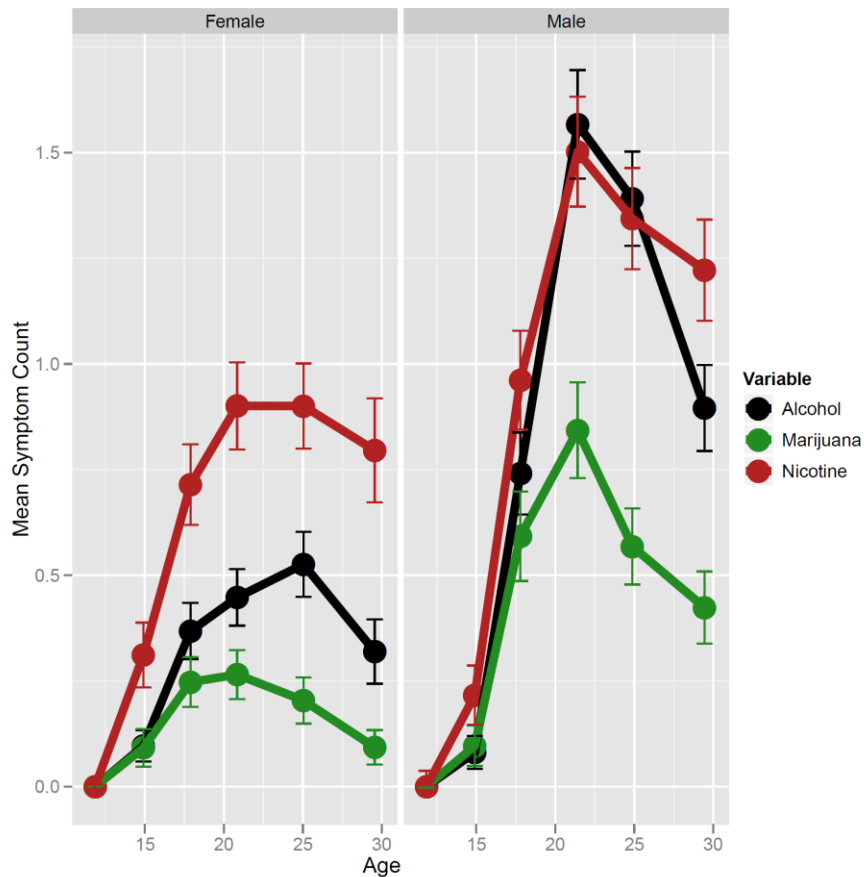**Figure 1. Mean Change in Symptom Counts with Age.**
Females are displayed on the left and males on the right. Each substance is plotted in a different color. Means for Nicotine Dependence, Alcohol Dependence, and Marijuana Dependence are in red, black, and green, respectively. Error bars represent 95% confidence intervals.

Cross-drug correlations decrease with age, as can be seen in the full bivariate

correlation matrix reported in Figure 2. For example, for males, the correlation between nicotine and alcohol declines from .61 at age 14 to .26 at age 29.

Females

	Nic (14)	Alc (14)	Mar (14)	Nic (17)	Alc (17)	Mar (17)	Nic (20)	Alc (20)	Mar (20)	Nic (24)	Alc (24)	Mar (24)	Nic (29)	Alc (29)	Mar (29)	
100	45	45	51	25	28	33	11	15	39	10	8				Nicotine — Age 14
61	100	35	31	25	19	21	12	15	25	14	9				Alcohol — Age 14
58	43	100	21	12	16	12	11	14	20	2	−1				Marijuana — Age 14
42	28	29	100	45	44	62	23	30	61	16	24	51	20	13	Nicotine — Age 17
29	28	26	53	100	41	34	37	26	35	25	29	27	19	10	Alcohol — Age 17
29	24	34	50	46	100	26	24	43	31	20	39	28	12	23	Marijuana — Age 17
28	20	19	57	39	33	100	29	36	74	24	21	66	26	16	Nicotine — Age 20
10	13	15	27	48	27	42	100	32	27	38	23	27	28	3	Alcohol — Age 20
22	16	28	35	40	49	46	46	100	34	27	54	34	24	33	Marijuana — Age 20
28	17	12	49	33	27	75	34	34	100	31	26	75	29	18	Nicotine — Age 24
10	14	13	23	36	19	29	51	31	32	100	29	31	36	12	Alcohol — Age 24
16	14	21	26	29	34	34	31	63	33	34	100	22	17	47	Marijuana — Age 24
33	21	17	45	27	25	67	26	28	77	23	29	100	34	25	Nicotine — Age 29
14	13	18	15	23	11	24	40	24	24	45	25	26	100	24	Alcohol — Age 29
21	23	25	28	26	35	31	27	55	31	25	67	30	30	100	Marijuana — Age 29

Males (lower triangle); columns by Age 14, Age 17, Age 20, Age 24, Age 29 (Nicotine, Alcohol, Marijuana each)

Figure 2. Within and Across Age Correlations among Substance Use Symptom Count Measures. Males are reported in the lower triangle and females in the upper triangle. Correlations (without decimals) are displayed within each colored box. To aid visualization, the matrix is a heat map, with hotter colors signifying higher correlations. The matrix is organized into blocks by age. Note the trend in the bolded diagonal blocks; the colors generally become cooler as one moves from the upper left to lower right, indicating a steady decrease in correlations among the substances over time. The off-diagonals have purposefully been partially obscured to focus the reader's attention on the block diagonal without omitting relevant information about the cross-age correlations. Females in the younger cohort had just begun their age-29 assessment, and thus the age-14/age-29 block is empty.

Factor models were fit to provide a formal test of changes in the cross-drug correlations.   Model fit in the full male sample was very good (χ^2=216.66, df=170, p=.009; ΔAkaike Information Criterion = 123.33; Root Mean Square Error of Approximation = .02). Model fit in the full female sample was also good (χ^2=432.32, df=170, p<.001; ΔAkaike Information Criterion = -92.32; Root Mean Square Error of

16

Approximation = .05). Model fit was good in both the male (χ^2=137.45, df=170, p=.97;

ΔAkaike Information Criterion = 202.55; Root Mean Square Error of Approximation <

.01) and female (χ^2=133.42, df=106, p=.04; ΔAkaike Information Criterion = 78.6; Root

Mean Square Error of Approximation = .04) early-use subsamples.

Table 2. Standardized Factor Loadings and Factor Variance Components for Each Age of Assessment in the Four Samples.

| | Factor Loadings | | | | | | | | Factor Variance ACE Components | | | | | |
| | Nicotine | | Alcohol | | Marijuana | | Mean Squared Loading | | A | | C | | E | |
	Est.	95% c.i.	Est.	95% c.i.	Est.	95% c.i.	Est.	95% c.i.	Est.	95% c.i.	Est.	95% c.i.	Est.	95% c.i.
Full Female Sample (N=1907)														
Age 14	.65	(.59,.70)	.71	(.66,.83)	.79	(.74,.84)	.52	(.48,.56)	-	-	-	-	-	-
Age 17	.60	(.55,.64)	.78	(.74,.81)	.73	(.68,.83)	.51	(.47,.54)	.72	(.56,.88)	.14	(.27,.29)	.14	(.10,.19)
Age 20	.51	(.45,.58)	.57	(.50,.64)	.60	(.54,.65)	.32	(.28,.36)	.42	(.14,.74)	.21	(0,.47)	.37	(.25,.49)
Age 24	.60	(.53,.60)	.54	(.46,.61)	.39	(.31,.45)	.27	(.24,.31)	.61	(.31,.61)	.09	(0,.34)	.30	(.19,.42)
Age 29	.69	(.60,.79)	.53	(.43,.62)	.33	(.25,.41)	.29	(.24,.33)	.36	(.09,.61)	.12	(0,.39)	.53	(.36,.68)
Full Male Sample (N=1775)														
Age 14	.74	(.69,.77)	.76	(.71,.80)	.77	(.73,.81)	.57	(.54,.61)	.35	(.16,.57)	.46	(.26,.63)	.19	(.14,.24)
Age 17	.70	(.65,.74)	.79	(.75,.82)	.72	(.67,.76)	.54	(.51,.57)	.69	(.50,.88)	.18	(.03,.36)	.14	(.10,.18)
Age 20	.65	(.60,.70)	.67	(.62,.72)	.70	(.64,.75)	.45	(.42,.49)	.58	(.37,.73)	.13	(.01,.32)	.29	(.22,.38)
Age 24	.55	(.48,.61)	.56	(.50,.62)	.56	(.49,.62)	.31	(.27,.35)	.54	(.30,.73)	.16	(.01,.39)	.30	(.19,.42)
Age 29	.47	(.41,.54)	.62	(.55,.69)	.59	(.52,.65)	.32	(.28,.36)	.47	(.47,.63)	.05	(0,.26)	.48	(.34,.61)
Females with at least 1 Symptom by Age 17 (N=486)														
Age 14	.49	(.35,.61)	.64	(.52,.74)	.82	(.69,.93)	.44	(.36,.51)	-	-	-	-	-	-
Age 17	.32	(.21,.42)	.58	(.46,.69)	.77	(.65,.89)	.34	(.29,.40)	-	-	-	-	-	-
Age 20	.23	(.12,.35)	.52	(.32,.73)	.68	(.49,.91)	.26	(.20,.34)	-	-	-	-	-	-
Age 24	.31	(.17,.44)	.63	(.48,.76)	.43	(.28,.59)	.22	(.16,.29)	-	-	-	-	-	-
Age 29	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Males with at least 1 Symptom by Age 17 (N=580)														
Age 14	.69	(.61,.76)	.73	(.64,.80)	.76	(.68,.83)	.53	(.47,.59)	-	-	-	-	-	-
Age 17	.47	(.37,.56)	.62	(.52,.72)	.65	(.55,.76)	.35	(.29,.41)	-	-	-	-	-	-
Age 20	.49	(.39,.59)	.57	(.46,.67)	.68	(.55,.81)	.35	(.28,.41)	-	-	-	-	-	-
Age 24	.40	(.28,.52)	.49	(.36,.62)	.48	(.34,.62)	.21	(.15,.28)	-	-	-	-	-	-
Age 29	.43	(.21,.44)	.61	(.48,.75)	.64	(.50,.79)	.30	(.24,.36)	-	-	-	-	-	-

"Est." is the maximum likelihood estimate and "95% c.i." the 95% confidence interval around that estimate. The "Average Squared Loading" is simply the average of the squared loadings listed for that age. Age-29 data is not available for the female subsample because the assessment is ongoing. A refers to the

additive genetic variance component of the factors, C to the shared environmental component, and E to the non-shared environmental component. A, C, and E estimates for the 14-year-old samples were poorly estimated due to lack of cross-twin co-variance in the symptom counts at that age (see Results), and are not provided for the females. A, C, and E estimates are not provided for the subsamples of early onset users because these samples are, by definition, within-individual and exclude co-twins who did not exhibit symptoms at the age-17 assessment. Resulting ACE estimates are therefore difficult to interpret.

For each model, the size of the standardized factor loadings generally decreased with age, declining from .6-.8 at age 14 to .3-.6 at age 29, as reported in Table 2. The table also shows (as do the grey lines in Figure 3) that the mean squared loading, which served as our metric of change, declined with age. Recall that the mean squared loading provides the average variance in the symptoms accounted for by the general factor. To illustrate how these estimates are calculated, we enter the loadings from Table 2 into Equation 1 for the full female sample. The mean squared loading was $(.65^2+.71^2+.79^2)/3=.52$ at age 14, and $(.69^2+.53^2+.33^2)/3=.29$ at age 29. To test if the age 14 to 29 decline was statistically significant, we constrained the mean squared loadings to be same at each age and examined decrement in model fit. The likelihood ratio test indicated a significant decline in the mean squared loading from age 14 to age 29 for all samples: full male sample ($\chi^2=158.4$, $df=4$, $p=3.2\times10^{-33}$), full female sample ($\chi^2=146.97$, $df=4$, $p=9.1\times10^{-31}$), male subsample of early users ($\chi^2=54.66$, $df=4$, $p=3.8\times10^{-11}$), and from age 14 to 24 for the female subsample of early users ($\chi^2=21.36$, $df=3$, $p=8.9\times10^{-5}$). We tested for sex differences by comparing the fit of an unconstrained model to that of a model that constrained the mean squared loading to be equal across males and females. Relative to males, females showed an earlier decline in comorbidity at age 20 ($\chi^2=24.18$, $df=1$, $p=8.8\times10^{-7}$), but at no other ages.
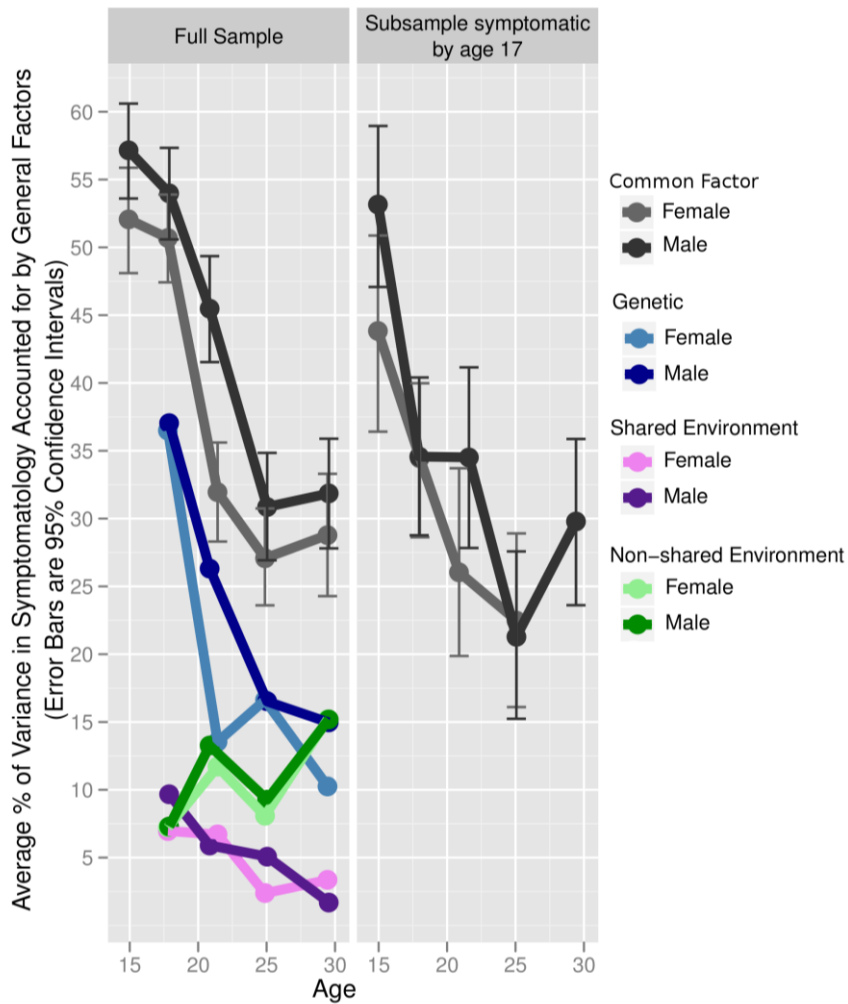
Figure 3. Percent of Symptom Count Variance Accounted for the General Factor at Each Age.
The grey lines in the figure show the decline in correlations over time as expressed by decline in the average percentage of variance accounted for by a general factor at each time point (i.e., the "Mean Squared Loading" column from Table 3). The full sample is shown on the left while the subsample of individuals who had at least one nicotine, alcohol, or marijuana symptom by their age-17 assessment is shown on the right. The decline for both sexes in both samples was statistically significant (see Results). Sample sizes are given in the text as well as in Table 2. In the full male and female samples the blue, green, and purple lines represent the proportions of the general factor phenotypic variance (grey lines) that are due to genetic (blue), shared environmental (purple), and non-shared environmental (green) variance. Male values are always represented by the darker hue. These estimates can be computed directly from values provided in Table 3 by multiplying the mean squared loading by the corresponding ACE value (e.g., by the A value to obtain the additive genetic variance which is plotted in blue). The majority of phenotypic decline in the full sample is due to a statistically significant decline in heritability, as well as a non-statistically significant decline in shared environmental variance. In contrast, non-shared environmental variance significantly increases with age. Note that age-14 estimates are not given because girls at this age had few symptoms (see Results). Genetic and environmental components are not given for the subsample as it was composed of selected individuals, and not selected twin pairs (see Methods).

Table 2 provides the A, C, and E components of each age's general factor, for each sample. While these estimates are useful, they must be multiplied by their respective factor loadings to determine the genetic and environmental impact on the individual symptom counts. To do this, we replaced the Var(Factor) term in Equation 1 with the A, C, and E variance components. The resulting values are displayed for each age in the left panel of Figure 3. Note that the scaling is such that adding the A (blue), C (purple), and E (green) components reproduce the mean squared loading (grey; also listed in Table 2). For example, at age 29 for males the phenotypic variance accounted for by the general factor was 32%, which was composed of 15% ($.32\times.47\times100\%$) additive genetic variance, 15% ($.32\times.48\times100\%$) non-shared environment, and 2% ($.32\times.05\times100\%$) shared environment ($15\%+15\%+2\%=32\%$). Unfortunately, symptoms were expressed relatively infrequently by both members of a female twin pair at age 14, rendering the age-14 additive genetic, shared environmental, and non-shared environmental components poorly estimated and hence excluded from Figure 3 along with the male estimates for consistency.

We tested for change in heritability, shared environment, and non-shared environment estimates in Figure 3 by fixing relevant parameters to be equal at ages 17, 20, 24, and 29, testing for decrements in model fit. For example, to test for a change in heritability, we fixed the A component of each general factor to be equal, and tested the fit of this model against the original model where all A components are freely estimated. The decline in heritability noticed in Figure 3 was significant for males (χ^2=23.86, df=3, p=2.7×10^{-5}) and females (χ^2=26.6, df=3, p=7.0×10^{-6}). It appears thus that the vast

majority of the phenotypic decline in Figure 3 is due to a decline in genetic variance. The increase in non-shared environment was also significant for males (χ^2=15.73, df=3, p=.001) and females (χ^2=11.48, df=3, p=.009). Changes in shared environment were not significant in either sex.

Discussion

We examined changes in the correlations among symptoms of nicotine, alcohol, and marijuana abuse and dependence from age 11 to 29. As seen in Figure 3, males and females show significant declines in these correlations from adolescence to adulthood. This was also true for an early-onset subsample of individuals with at least one symptom of nicotine, alcohol, or marijuana dependence by age 17. Correlation among disorders or their symptoms is evidence that those disorders share etiology. The results suggest that the shared etiology contributing to nicotine, alcohol, and marijuana dependence symptoms diminish over time. That is, younger individuals tended to use these three substances indiscriminately, whereas older individuals began to show a preference for one substance over others. Despite declines in the correlations, the rates of use (Figure 1) continued to climb throughout late adolescence and early adulthood. Finally, after age 17 the correlations among symptoms became less attributable to pleiotropic genetic effects and increasingly a result of non-shared environmental influences (Figure 3), indicating that the types of etiological processes contributing to the variation in use of multiple drugs is gradually changing during the transition to adulthood.

Several processes might contribute to the transition from general to specific influences. For one, adolescents are more impulsive and risk-taking than adults

(Steinberg, Cauffman, Woolard, Graham, & Banich, 2009). Personality traits such as disinhibition and sensation-seeking are not predispositions to use any particular substance, but rather to use whatever substances might be available (Bechara, 2005; Iacono, et al., 2008). Additionally, neuro-developmental changes relevant to behavioral disinhibition continue throughout adolescence. For example, by adolescence, the nucleus accumbens—important in reward sensitivity—is well-developed, but poorly regulated by a still maturing prefrontal cortex (Casey, Getz, & Galvan, 2008; Casey, Tottenham, Liston, & Durston, 2005), resulting in deficits of top-down control over the reward system that slowly improves into early adulthood. This developmental window is the same time period that we observed decreases in the comorbidity among different substance use disorders, suggesting that disinhibitory mechanisms with known neurological substrates may be in play. Further supporting this hypothesis, the earlier decline in comorbidity for females relative to males from age 17 to 20 (Figure 3) is consistent with the earlier pubertal (Tanner, Whitehouse, & Takaishi, 1966), cortical (Lenroot et al., 2007), and personality (Klimstra, Hale, Raaijmakers, Branje, & Meeus, 2009) maturation in females compared to males.

Our results are consistent with individual differences in drug reinforcement. That is, initial drug use that begins in adolescence is characterized by relatively indiscriminant experimentation. Due to individual differences in the reinforcing effects of different drugs, however, people may tend eventually to restrict their use to those drugs that provide the greatest reinforcement. There are myriad etiological mechanisms relevant to individual differences in drug reinforcement, including drug metabolism effects, drug

24

availability, social rewards and punishments, and differences in drug-specific

neurological sensitivity that may further be moderated by drug and alcohol neurotoxic

mechanisms (Bechara, 2005; Goldstein & Volkow, 2002). Future research might

consider these covariates in characterizing the transition to substance specialization.

While consistent with the theory that disinhibition accounts for comorbidity

among substance use disorders, the results are not entirely inconsistent with the gateway

hypothesis (Kandel & Jessor, 2002). The gateway hypothesis holds that using one drug

leads to using other drugs, perhaps due to the experienced high and an increased desire to

obtain bigger and better highs. If correct, the theory would predict the opposite pattern of

comorbidity we observed; that is, correlations should be low at younger ages and increase

over time. At younger ages, most people would not yet have used their first or second

gateway drug, and would not have had time to explore other drugs. Over time, the

correlations among drugs should increase as the initial drug use would cause them to use

other drugs. In fact, we found that as drug dependence symptoms increased (Figure 1),

the magnitude of the associations among different drugs declined (Figure 3); opposite the

pattern predicted by the gateway theory. This conclusion is consistent with a growing

body of literature inconsistent with parts of the gateway theory (Iacono, et al., 2008;

Irons, McGue, Iacono, & Oetting, 2007; Kendler, Jacobson, et al., 2003; Tarter,

Vanyukov, Kirisci, Reynolds, & Clark, 2006; Vanyukov et al., 2003). That said, the

increase in non-shared environmental effects observed in Figure 3 could contain etiology

analogous to a gateway process, in that there are environments experienced by an

individual that contribute to dependence to multiple drugs in that person, but not that

person's co-twin. While the range of possible environments is vast and the effect is small, it could very well include drug use (e.g., impaired cognition caused by neurotoxicity) and/or other risk factors for non-specific drug use such as occupational, social, and legal problems.

Limitations

While a strength, the use of a community-representative sample may not apply to clinical populations. The fact that results were consistent for the subsample of individuals who were symptomatic at age 17 suggests this is not a major concern. We used symptom counts because symptoms are relevant to the DSM clinical literature and provide measurement consistency from age 11 to 29. In supplementary analyses we evaluated other measures of quantity and frequency of nicotine, alcohol, and marijuana use (not shown). These measures had higher means and variances at younger ages, and resulted in the same trends as presented here for symptoms.

In the United States adolescent development is confounded with substance-use-relevant environmental changes. Adolescents gradually experience increased autonomy and financial freedom from caregivers. The purchase of tobacco and alcohol becomes legal at age 18 and 21, respectively, while marijuana use is always illegal. Undoubtedly, these environmental influences impacted symptom means and correlations. However, one cannot disentangle these influences from other behavioral and neurological changes, at least within a single culture. Cross-cultural and cross-generational studies, for example comparisons between societies that differ in drug laws, are required to unravel maturational and environmental changes during development.

Study 2. Genome-wide SNP Scores Predict Behavioral Disinhibition,

Alcohol Use, Drug Use, and Nicotine Use, as well as their relationships

with one another.

Disinhibition is a behavioral trait hypothesized to represent a general vulnerability in the development of substance use disorders (Iacono, et al., 2008; Zucker, Heitzeg, & Nigg, 2011). Those with greater levels of disinhibition are thought to act more impulsively, have greater thresholds for thrill-seeking, and not consider as deeply the long-term consequences of their actions. These individuals are considered more likely to use substances and have a more difficult time quitting. Evidence for this hypothesis comes from a variety of research designs summarized in (Iacono, et al., 2008). There is additional evidence that the relationships between substance use measures and other disinhibitory measures are moderately to highly heritable (Krueger, et al., 2002; Vrieze, Iacono, & McGue, 2012), and that genes account for a large proportion of covariance among substance use disorders (Kendler, Jacobson, et al., 2003; Kendler, Prescott, et al., 2003). However, research on measured genetic variants, such as single nucleotide polymorphisms (SNPs), has not been successful in locating individual genes or genetic variants responsible for the genetic variance in common substance use disorders, although there are notable exceptions (Furberg et al., 2010; Luczak, Glatt, & Wall, 2006; Schumann et al., 2011).

It appears that for many complex traits the effects of individual genetic variants are small (Manolio et al., 2009). Millions of SNPs have been the most commonly studied

genetic variants, and massive sample sizes have been required to reliably separate the genetic signal from noise. The problem is two-fold. First, individual SNPs have small effects (e.g., account for < 0.5% of phenotypic variance). Second, the atheoretical approach of genome-wide association studies involves the simultaneous testing of 500,000 to 5 million SNPs, and incurs a very substantial multiple testing burden (Hirschhorn & Daly, 2005).

One effective solution has been to assemble mega-samples of hundreds of thousands of individuals to obtain sufficient statistical power to detect these small effects. Such endeavors have identified hundreds of variants (Visscher, Brown, McCarthy, & Yang, 2012) for complex traits like height (Allen et al., 2010), BMI (Speliotes et al., 2010), and lipid levels (Teslovich et al., 2010), as well as for complex diseases such as Crohn's (Franke et al., 2010) and Type-2 Diabetes (Voight et al., 2010). These studies are highly informative both in identifying relevant genome-wide significant (i.e., $p < 5e^{-8}$) SNPs and candidate genes (e.g., see Teslovich, et al., 2010, for a nice example) and they also provide information on the distribution of SNP associations with phenotypes.

The GWAS meta-analyses on height and BMI, for example, provide genome-wide scoring results, in addition to reporting individual genome-wide significant SNPs. Genome-wide scoring is a process by which SNPs are weighted and combined to produce, for each person, a single aggregate SNP score, and the phenotype is then regressed on that score. SNPs, or any predictor for that matter, can be weighted in many ways. In practice, each SNP is often weighted according to the univariate analysis of the phenotype regressed on that SNP, controlling for covariates.

The present study uses genome-wide scoring in a moderately-sized twin and

adoptive family study sample (N = 7188) to investigate the genetic architecture of several

measures of substance use pathology that have been described in detail previously

(Hicks, Schalet, Malone, Iacono, & McGue, 2011), and includes measures of nicotine

use, alcohol consumption, alcohol dependence, illicit drug use, and behavioral

disinhibition.

The present sample contains a large number of twins, and genome-wide scoring

results from the present sample are compared to heritability estimates obtained using

standard twin methodology, as well as heritability estimates based on common SNPs

using Genome-wide Complex Trait Analysis (Yang, Lee, Goddard, & Visscher, 2011;

Yang et al., 2011). The results inform the extent and form of polygenetic heritability for

substance use traits, and provide guidance for future study of behavioral disinhibition and

substance use traits and disease.

Method

The sample used in this research has been described in detail elsewhere (Iacono &

McGue, 2002; Miller et al., submitted). In short, it is composed of two studies of

Minnesota families: a community-representative sample of twins and their parents, as

well as a study of adoptive families. Sample sizes for the twin and adoptive families are

provided in Table 1. Twin families are further divided into three cohorts: (A) a

longitudinal sample (N = 1139) of 17-year-old twins (N = 1139) and 11-year-old twins

(N = 1167) first assessed between 1989-1996, followed regularly at ages 20, 24, and 29;

(B) a longitudinal sample (N = 1167) of 11-year-old twins assessed at ages 11, 14, 17, 20,

29

24, and 29, with their age-17 assessment occurring sometime between 1996-2003; and

(C) another longitudinal sample (N = 571) of 11-year-old twins assessed at ages 11, 14,

and 17, with their age-17 assessment occurring between 2005 and 2010. Parents were

typically assessed at intake, regardless of cohort. In total, the present work was on 7188

Caucasian participants from 2300 families.

Phenotypic Measures

Development and construct validity of the phenotypes used in this study have

been described extensively in a development report (Hicks, et al., 2011). Nicotine Use

and Dependence (NIC) was composed of lifetime measures of DSM-3R symptoms of

nicotine dependence, as well as frequency and quantity of nicotine use during the period

of an individual's heaviest use. Alcohol Consumption (CON) was composed of lifetime

intoxications, lifetime maximum number of drinks in a 24-hour period, and frequency of

alcohol use (e.g., number of days per week drinking at least one beverage). Alcohol

Abuse/Dependence (DEP) was composed of DSM-III-R, DSM-III, Research Diagnostic

Criteria, and Feighner Criteria, and included symptoms and signs of alcohol dependence

and abuse such as tolerance and withdrawal, social and occupational problems due to

drinking, and compulsive drinking (e.g., little time for anything but drinking). Drug Use

(DRG) was composed of a measure of the number of times an individual had ever tried

marijuana, as well as a count of the number of classes of other illicit drugs they had ever

tried (e.g., stimulants, hallucinogens, PCP, etc.), as well as symptom counts for the drug

to which they reported the most symptoms. Finally, Behavioral Disinhibition (BD) was

composed of measures of DSM-IIIR Conduct Disorder, Adult Antisocial Behavior, and other antisocial and non-normative behavior such as early sexual intercourse.

Genotyping

Details of the genotyping procedures are provided in (Miller, et al., submitted). In short, genome-wide genotyping was done on the Illumina Human660W-Quad Array, which contains a total of 561,490 SNPs. A total of 32,153, or 5.7% of the markers attempted, failed one or more of these quality control filters, leaving 527,829 markers that passed all QC filters. Genotyping was attempted on samples from 7,438 participants. A total of 160 (2.2%) of samples failed quality control filters and were dropped from the present analysis. Only one MZ twin from each MZ twin pair was genotyped. The genotype was then wholly imputed to the MZ cotwin (N = 1127) for a final GWAS sample of 8405 individuals.

The majority of sample self-identified as White (90.4%), but we selected individuals for the current analysis on the basis of genetic principal components with Eigenstrat (Price et al., 2006). Since Eigenstrat is sensitive to close relatives, one member of each close relative pair was excluded. Full details are provided in (Miller, et al., submitted). To identify individuals as white for the present analysis we first found the centroid in the 10-dimensional principal component space, and then computed the distance of every individual from this centroid. A hypersphere surrounding all individuals was then contracted until further contraction resulted in shedding a large proportion of self-reported white subjects relative to non-white subjects. The process resulted in a sample of 7702 of putatively white individuals, including 101 for whom we

did not have self-reported ethnicity and 46 who had originally self-reported as something other than white. The ten principal components were also used as covariates in all analyses to correct for any spurious effects arising from population stratification.

Heritability Estimates with Biometry and Genome-wide Complex Trait Analysis

The present sample allows several methods to estimate heritability of the phenotypes. First, a large portion of the sample is composed of twins. We used knowledge of twin zygosity and standard biometric statistical models to compute the additive genetic variance-covariance matrix (A), shared environmental variance-covariance matrix (C), and non-shared environmental variance-covariance matrix (E) components of the 5x5 variance-covariance matrix of our five measures. This is the standard multivariate ACE model (Neale & Cardon, 1992). Variance-covariance component estimates were estimated by full information maximum likelihood after correcting for fixed effects of sex, age, year of birth, and the first 10 genetic principle components computed from Eigenstrat. Model fit was evaluated with accepted indexes of fit, including a likelihood ratio test and the Akaike Information Criterion (AIC). The likelihood ratio test is sensitive to sample size and correlational magnitude, and so is often augmented with measures of fit like the AIC, which have attractive theoretical properties not shared by the likelihood ratio, such as minimization of mean squared error of estimation (Vrieze, 2012).

A second way to estimate heritability is to consider the additive effect of all SNPs considered simultaneously using Genome-wide Complex Trait Analysis (GCTA). GCTA has become increasingly used to provide an estimate of the heritability in a trait due to

measured SNPs (Yang, Lee, et al., 2011; Yang, Manolio, et al., 2011). The method considers the joint effect of all SNPs considered simultaneously as a random effect, and estimates the variance in the phenotype attributable to this random effect. In practice, the method computes the genetic relatedness based on SNPs between all pairs of individuals in the sample, this genetic relatedness matrix (GRM) is then used as input in the random effects model, and the similarity among individuals in genetic relatedness predicts the similarity in phenotypic relatedness. In a sample of unrelated subjects the method produces the variance in the trait accounted for by the SNPs, because the relationships between genetically unrelated subjects are not influenced by shared environment or non-SNP genetic variance. In a sample of related subjects, such as the families used in the present study, phenotypic relatedness and genetic relatedness are confounded in important ways that must be addressed. If estimates are based on everyone, then SNP-based genetic relatedness and phenotypic relatedness are confounded with rare and non-additive genetic relationships (e.g., monozygotic twins share almost all variants, including rare and common SNPs) and shared environmental effects due to shared family experiences.

To account for familial confounding, we used GCTA on four samples. (1) The best way to estimate the random effect of SNPs (the aggregate effect of common SNPs on the phenotype) is with a large sample of unrelated individuals (Yang et al., 2010). The largest such sample in the present study consists of all genetically unrelated parents (N = 3542). To determine genetic relatedness in this parent sample we excluded one individual of every pair of individuals who had a genetic relatedness of $\geq$ .025 as

33

calculated by the genetic relatedness matrix produced by GCTA on the full sample. (2) To help inform the biometric twin heritability estimates, we also estimated the random effect of SNPs on an unrelated sample of the youth offspring (N = 1784), including both the twins and non-twin siblings. Ideally, this would provide an estimate of the aggregate effect of common SNPs in the offspring youth sample and would be comparable to that produced by the unrelated parent sample. (3) We conducted the same analysis on the full sample of youth offspring (N = 3336), without concern for genetic relatedness. Because this analysis confounds phenotypes, genotypes, and shared environment, it should return a genetic random effect equivalent to the sum of genetic and shared environment from the biometric analysis (i.e., A + C). (4) Finally, we estimate the random genetic effect in the full sample (N = 7188), which should provide an estimate of the random effect somewhere between the unrelated sample and the youth offspring sample, as the full sample has less of a shared environment confound than the twin sample (i.e., parents are phenotypically related due to shared environment, but are mostly not genetically related).

Genome-wide Scoring Procedure
       The biometric and GCTA methods provide variance component estimates of the aggregate effect of genetic variants. They are limited in that they do not provide weightings for individual SNPs, nor can they be applied to new samples in attempts to predict genetic loading for some trait. Genome-wide scoring, on the other hand, does return this information.

Scoring proceeded in a series of steps. First, the phenotype was residualized using a linear regression on covariates of sex, generational status (parent or child), age, year of birth, and the first 10 principal components produced by Eigenstrat. A GWAS was then conducted on the residualized phenotype, producing a univariate regression weight for the minor allele count for each SNP. Minor allele counts for each SNP were then multiplied by their corresponding regression weight and summed to form a single score for each participant in the sample. This sumscore was then validated by correlating it with the residualized phenotype. Squaring the correlation gives the variance in the phenotype accounted for by the SNP score.

Gross overfitting is expected when the same sample is used to generate and validate the SNP score, especially when the number of predictors is much greater than the number of subjects ($p >> N$). To control for overfitting we employed a k-fold cross-validation technique (Breiman & Spector, 1992; Hastie, Tibshirani, & Friedman, 2009). For this study we set the number of folds k to be 10. To accomplish this, subjects were split into 10 roughly equal subsamples. The scoring algorithm described above is conducted by combining 9 subsamples, providing a set of SNP weights based on the 9 subsamples combined. These weights were then applied to the minor allele counts in the 10th sample and correlated with the phenotype in that sample, producing an unbiased estimate of the cross-validated validity of the SNP score. This same procedure is used for every combination of the 10 samples, such that every single subject is in a development sample nine times and in the test sample once.

Because this sample is composed of families, and individuals within families are correlated with respect to genotypes and phenotypes, we always kept individuals from the same family within the same subsample. This prevented the algorithm, for example, from deriving the SNP score on one twin and cross-validating it on the other – clearly in that case we expect prediction bias given correlation between twins on the phenotype and the genotype.

SNPs were also filtered on the basis of linkage disequilibrium (LD). If two tag SNPs are in LD with a causal variant, and in LD with each other, then both SNPs will show a relationship with the phenotype, despite the fact that the two SNPs are redundant. To avoid overcounting such redundant SNPs prior studies have imposed strict LD cutoffs, such that no two SNPs included in the set of prediction SNPs can have LD $r^2 >$ .05 (e.g., Allen, et al., 2010). We chose to evaluate three different LD cutoffs: $r^2 = .05$, .50, and 1.0 (i.e., no cutoff). The cutoff was imposed in the following way. First, all SNPs were regressed on the phenotype and their univariate weights and p-values recorded. SNPs were then sorted according to p-value. The most significant SNP was selected, and all SNPs with LD greater than the cutoff were culled from the list of SNPs. Then the remaining second-most significant SNP was considered and all SNPs in LD with it were removed. This process was completed until the least significant SNP was considered. *p*-value thresholds for significance were varied from .0001 (allowing very few SNPs in the score) to 1.0 (allowing all SNPs).

Finally, to increase confidence in the scoring results, we simulated three types of phenotypes. First, we simulated a phenotype with no genetic association, which we refer

to as "Random." Second, we simulated phenotypes from 10,000, 50,000, and 100,000

causal SNPs under an additive model with normally distributed regression coefficients.

Third, we simulated phenotypes from 10,000, 50,000, and 100,000 SNPs under an

additive model with uniformly distributed regression coefficients. The phenotype in both

the normal and uniform scenarios was simulated to be 20% heritable, in line with

expectations based on the GCTA analyses reported in the current study.

Genome-wide scoring with 10-fold cross-validation is computationally

demanding. The computational demands prevented us from conducting permutation or

other tests of statistical significance. Fortunately, the cross-validation statistic in use here

is the Pearson correlation and is amenable to short-hand tests of significance. The

standard error of the Pearson after z-transformation is $1/\sqrt{(N-3)}$, and $z = \arctan(r)$. A

significant t-score $= 1.96$. The p-value for z and any N is thus approximately $\Phi(z \times \sqrt{N})$,

where Φ is the probability density function of the standard normal distribution. The

average within-family correlation, averaging over all five phenotypes, was .24.

Multiplying the total sample size by one minus the squared average within family

correlation yields $7188\times(1-.24^2)\approx6774$, an estimate of the effective sample size. When N

$= 6774$, a correlation coefficient r must be greater than .02 to be significant at p < .05. If

we are conservative, and set our effective sample size at 5,000 individuals, then a

correlation coefficient must be $r > .024$ to be significant at $p < .05$.

For all analyses we covaried out the linear effects of age, sex, year of birth,

generational status (parent/offspring), and the first 10 genetic principal components.

The *k*-fold cross-validation algorithm was programmed in the *R* Environment

2.10.1 (R Development Core Team, 2011), and GWAS conducted using the GenABEL

package 1.6-9 (Aulchenko, Ripke, Isaacs, & van Duijn, 2007). Biometric twin models

were estimated with the OpenMx package 1.2 (Boker, et al., 2011). GCTA analysis used

the GCTA program 0.93.9 (Yang, Lee, et al., 2011). Scripts available upon request.

Results

Descriptive statistics for the substance use phenotypes and behavioral

disinhibition are provided in Table 3. Means and Variances of the Five Substance Use

Phenotypes.. Figure 4 provides the biometric estimates, based on the twin sample alone

(average age = 17), of the heritable, shared environmental, and nonshared environmental

components of the correlation matrix among the five phenotypes. The variance in each

phenotype, as well as the covariances among phenotypes, are largely due to heritable

variation. However, there are also significant shared environmental effects, both on the

variances and covariances. Finally, there are significant non-shared environmental

effects, especially on the variances. Fit statistics for the ACE, ADE, and AE models are

listed in Table 4. Fit of the Biometric Models. The ACE model fit best, followed by the

AE and ADE models, according both to likelihood ratio tests and the AIC.

38
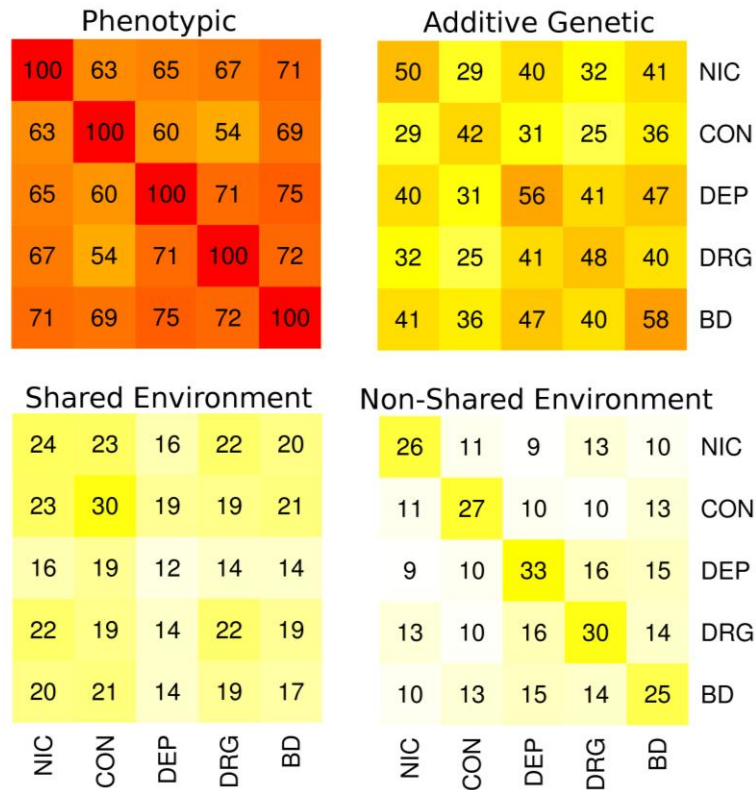
Figure 4. Phenotypic correlations and biometric decomposition.
Shown here are the phenotypic correlation matrix, as well as the additive genetic, shared environmental, and non-shared environmental component matrices. The component matrices are scaled such that they sum elementwise to produce the full phenotypic matrix. All entries are significant at $p < .05$. Estimates are based solely on the twins, who have an average age of 17 years. NIC = Nicotine Use; CON = Alcohol Consumption; DEP = Alcohol Dependence; DRG = Drug Dependence; BD = Behavioral Disinhibition.

Table 3. Means and Variances of the Five Substance Use Phenotypes.

Sample	Sex	N	Age M (SD)	Means					Variances				
				NIC	CON	DEP	DRG	BD	NIC	CON	DEP	DRUG	BD
Twin Families													
Parents	Male	1396	43.87 (5.65)	.53	.80	.86	.41	.63	1.16	.42	1.64	1.36	1.10
	Female	1587	41.67 (5.15)	.18	.09	-.06	-.01	-.33	1.11	.35	.63	.83	.55
MZ Twins	Male	873	17.80 (0.61)	-.21	-.23	-.25	-.17	.24	.70	1.38	.63	.90	1.17
	Female	990	17.88 (0.77)	-.44	-.64	-.51	-.39	-.44	.53	1.0	.25	.47	.66
DZ Twins	Male	457	17.75 (0.49)	-.22	-.35	-.26	-.18	.25	.70	1.45	.61	.92	1.21
	Female	557	18.03 (0.81)	-.32	-.49	-.40	-.28	-.26	.65	1.05	.46	.65	.82
Adoptive Families													
Parents	Male	384	48.43 (4.46)	.22	.49	.44	.45	.22	1.13	.49	1.21	1.35	.73
	Females	485	46.53 (4.11)	-.18	-.11	-.21	.04	-.50	.91	.32	.47	.78	.28
Bio Sibs	Male	168	18.16 (1.01)	-.26	-.14	-.09	.15	.35	.65	1.52	.78	1.37	1.31
	Female	194	18.33 (1.14)	-.46	-.49	-.37	-.15	-.27	.46	1.05	.41	.72	.76
Adopted Sibs	Male	50	18.54 (0.87)	.09	.13	.05	.45	.68	.79	1.33	.98	1.47	1.20
	Female	47	18.37 (1.05)	-.24	-.25	-.29	.04	-.02	.51	.80	.61	.73	.59

Each variable in the full sample was scaled to have mean zero and variance one. NIC = Nicotine Use; CON = Alcohol Consumption; DEP= Alcohol Dependence; DRG = Drug Dependence; BD = Behavioral Disinhibition; BD = Behavioral Disinhibition. MZ stands for monozygotic twins and DZ stands for dizygotic twins.

Table 4. Fit of the Biometric Models

Model	-2LL	Free Parameters	χ^2(df), p-value	AIC
ACE	114287.4	50	N/A	85617.37
ADE	114365.3	50	N/A	85695.34
AE	114314.9	35	27.5(15), .02	85614.88

For the ACE, ADE, and AE models, the A refers to the additive genetic component, D to the dominant/recessive genetic component, C to the shared environmental component, and E to the non-shared environmental component. -2LL is minus to the log of the likelihood. AIC is the Akaike Information Criterion. The χ^2 is for comparison between the AE and the ACE model. The significant p-value of .02 indicates that the AE leads to decrement in fit.

GCTA results are given in Figure 5 and differ depending on the sample and the

GRM cutoff. First, the best estimate of the aggregate effects of the SNPs was produced

from the sample of genetically unrelated parents, as determined by using a cutoff of <

.025 on the genetic relatedness matrix. This estimate is "best" because it contains the

largest subset of genetically unrelated individuals that can be obtained in this sample, and

the resulting estimate will not be confounded due to non-common-SNP genetic effects

and/or strong shared environmental effects. As can be seen in Figure 5, the estimates of

phenotypic variance accounted for by the aggregated SNPs range from .16-.22. All

estimates were statistically significant except DEP (Alcohol Dependence). The full

sample estimates with no GRM cutoff yielded much higher estimates, consistent with the

notion that rare-SNP, non-additive, non-SNP, and/or shared environmental effects are

contributing to phenotypic similarity, sometimes substantially. It appears, however, that

the GCTA estimates from the full sample are highly similar to the additive genetic

estimates obtained in the biometric twin results, indicating only small inflation in the

GCTA results due to shared environmental confound in the full sample. The GCTA

results in the full sample are .59, .51, .51, .49, and .60 for NIC, CON, DEP, DRG, and

BD. The same biometric twin estimates are .50, .42, .56, .48, and .58.
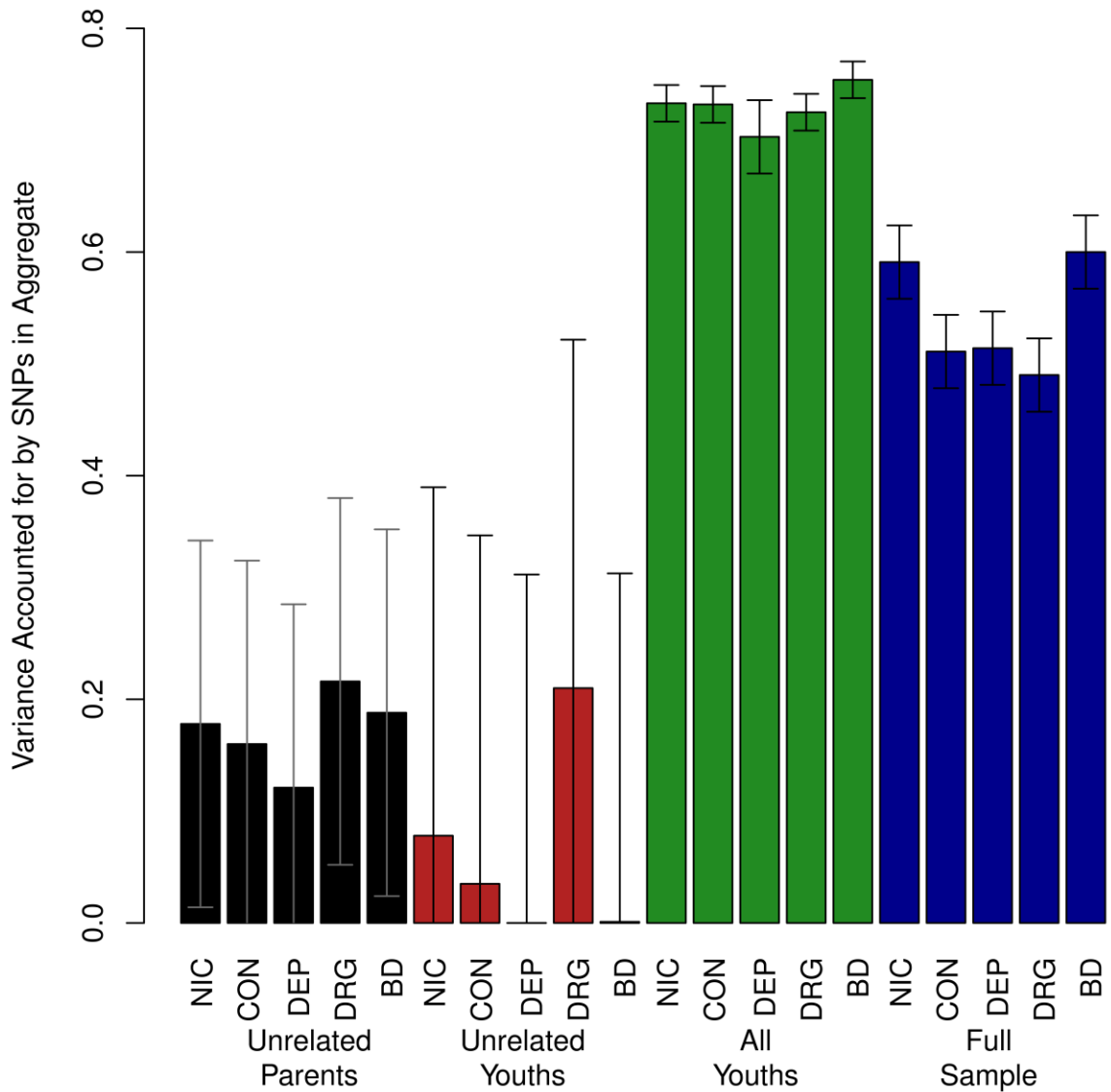
Figure 5. Genome-wide Complex Trait Analysis (GCTA).
Genome-wide Complex Trait Analysis (GCTA). The GCTA results are provided for each phenotype in a variety of samples. Unrelated individuals were defined as those having a genetic relatedness estimated by GCTA to be < .025 (more distantly related than third cousins). The samples are: A) All unrelated parents (N = 3542), B) unrelated youths (N = 1784), C) All youths (N = 3336), and D) the full sample (N = 7188).

In the sample of unrelated youths (GRM cutoff of .025) estimates of the aggregate

SNP effects are small and highly unstable, perhaps due to less phenotypic variability and

a relatively small sample (N = 1784). When evaluating the full sample of youths (N =

3336), which does not control for non-SNP genetic relatedness or shared environmental

effects in the youths, we see the estimates approach .70, perhaps indicating a stronger

role of shared environment in the youth-only sample versus the full sample of parents

and children. In fact, if we sum the heritabilities and shared environmental components

reported in Figure 4, we find that they are strikingly similar to the GCTA estimates on the

full set of youths. The GCTA values for the full sample of youths (in green in Figure 5)

are .73, .73, .70, .73, .75, for NIC, CON, DEP, DRG, and BD, respectively; Adding the

heritabilities to the shared environmental components from the twin results in Figure 4

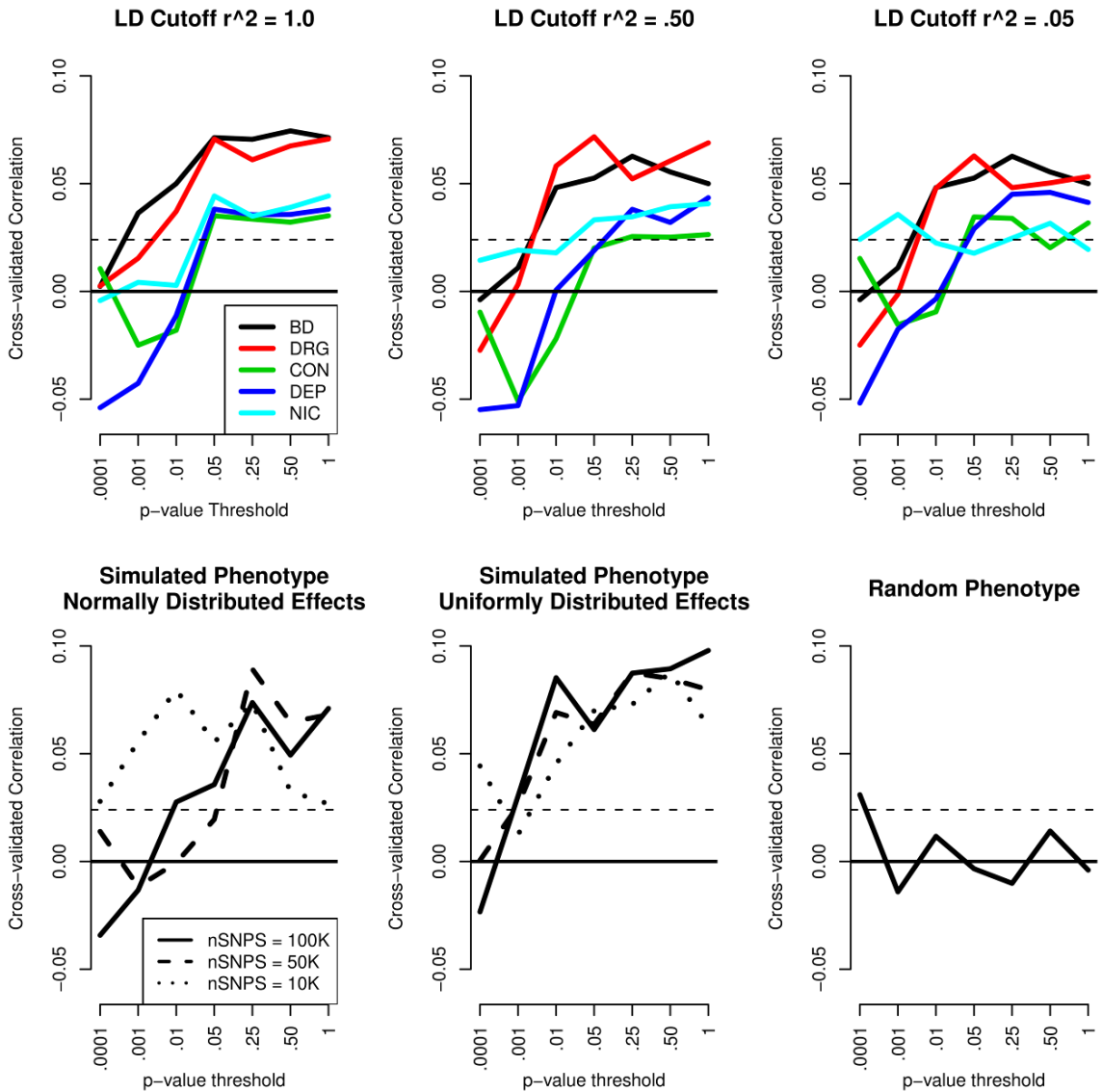gives almost the exact same estimates: .74, .72, .68, .70, and .75.

Figure 6. Cross-validated Genome-wide Scoring Results.
Cross-validated Genome-wide Scoring Results. The top panel of three graphs provides the empirical results for the four substance use phenotypes and behavioral disinhibition. Each graph provides the seven p-value thresholds under consideration. The three top graphs only differ in the LD cutoff imposed (1.0, .50, and .05). The bottom row provides results from three kinds of simulated phenotypes. First, a simulated phenotype with a normal distribution of SNP regression coefficients, for each of the 7 p-value thresholds and three different polygenetic scenarios (100,000, 50,000, and 10,000 associated SNPs). Second, the same scenario except with uniformly distributed effects. Both of these simulated phenotypes were simulated such that the SNPs in aggregate accounted for 20% of the variance in the phenotype. Third, a completely random phenotype with no SNP associations. The bold horizontal line in each graph is zero. The dotted line represents a correlation that would be significant at p < .05, conservatively assuming an effective sample size of 5000. BD = Behavioral Disinhibition; DRG = Drug Dependence; CON = Alcohol Consumption; DEP = Alcohol Dependence; NIC = Nicotine Dependence.

While moderately strong aggregate SNP effects were observed in the full sample, the genome-wide scoring procedure was unable to tap more than a small fraction of that variance. The results from the scoring procedure are given in Figure 6. The top row of sub-figures in Figure 6 provide the genome-wide scoring results for the five phenotypes, under seven p-value thresholds and three LD thresholds. To explain, consider the top right figure. Here we imposed an LD threshold of .05. That is, SNPs were excluded whenever they were in LD $> .05$ with a nearby, more significant, SNP. Each phenotype was then analyzed under seven p-value thresholds. Stringent p-value thresholds, such as excluding all SNPs with univariate effect significant at $p > .0001$ produced essentially null results for every phenotype but nicotine. As the p-value threshold was relaxed, improvement is seen for every phenotype (sans nicotine), until the effect plateaus at around $p = .05$. This pattern of results is generally true for each of the LD cutoffs (each of the three graphs in the top row of Figure 6). Most obviously, there appears to be a polygenetic effect—the variance accounted for in the phenotype increases substantially as the p-value threshold is relaxed and more SNPs are included in the scores. Results do not appear to be significantly or substantively affected by choice of LD cutoff. The polygenetic effect, and the pattern of results given different p-value thresholds, is true for the phenotypic data as well as the simulated phenotypes, regardless of the number of SNPs contributing to the phenotype in the simulations. The random, non-genetically-related phenotype (bottom right figure in Figure 6) shows no association, as expected.

Table 5 reports cross-validated correlations between genome-wide scores generated on one phenotype and correlated with a different phenotype. Correlations

46

among the drugs are of smaller, but similar, magnitude, ranging from 0.03 to 0.07, indicating small but detectable pleiotropic SNP effects in this sample.

Table 5. Cross-phenotypic correlations among genome-wide scores.

	NIC	CON	DEP	DRG	BD
NIC	.040				
CON	.030	.032			
DEP	.035	.043	.034		
DRG	.046	.043	.045	.062	
BD	.041	.053	.051	.068	.064

NIC = Nicotine Use; CON = Alcohol Consumption; DEP = Alcohol Dependence; DRG = Drug Use; BD = Behavioral Disinhibition. The diagonal represents the cross-validated correlation between the genome-wide SNP score and the actual phenotypic measurements. Off-diagonals represent the mean correlation between trait 1's genome-wide score and trait 2's phenotypic measurements, plus the correlation between trait 2's genome-wide score and trait 1's phenotypic measurements. Scores here are taken from the models with no LD cutoff and no p-value cutoff. All correlations are significant at $p < .05$, assuming an effective sample size of 5000, as discussed in the Method.

Discussion

The substance use phenotypes described here appear to be, like other complex traits and diseases, polygenetic and moderately to highly heritable. Heritabilities estimated from the subsample of twins ranged from 42% for Alcohol Consumption to 58% for Behavioral Disinhibition (Figure 4). Additive SNP effects estimated by GCTA on the parent sample account for 16% of the variance in Alcohol Dependence to 22% of the variance in Drug Use. While the aggregate additive SNP effect, estimated by GCTA is relatively large (e.g., 10-30% of total phenotypic variance; Figure 5), identifying and summing actual individual SNPs with genome-wide scoring yields much weaker effects, accounting for around 0.25% of the variance in the substance use and behavioral

disinhibition measures. Dividing the GCTA results estimated in the parent sample by the twin-estimated heritabilities allows us to estimate the total heritable variance accounted for by the aggregate SNP effect from GCTA. We estimate that the additive SNP effect accounts for 21% (Alcohol Dependence), 32% (Behavioral Disinhibition), 36% (Nicotine Use), 38% (Alcohol Consumption), and 45% (Drug Dependence) of the heritable variance in these traits.

There are apparent pleiotropic effects observed in the biometric twin heritability estimates. The genetic correlations between disorders (Figure 4) were relatively high (.25 - .47), such that one can expect genetic variants for one disorder to predict from 7% to 22% of the phenotypic variance in another disorder. A genome-wide score developed on one phenotype accounted for .09% to 0.46% of the variance in other phenotypes (squaring the correlations provided in Table 5), again indicating some extent of pleiotropy in the associated SNPs.

The disinhibitory hypothesis, that disinhibition represents a substantial source of general and genetically-based risk for substance use, was strongly supported by these results. First, that behavioral disinhibition shares genetic etiology with substance use disorders is supported by the biometric twin results given in Figure 4, in that the measure of behavioral disinhibition was highly genetically correlated with the other traits. What is more, the genome-wide scores generated on behavioral disinhibition were predictive of all the substance use traits (Table 5), indicating the existence of a polygenetic SNP-based relationship between disinhibition and substance use, but limitations in statistical power prohibit stronger conclusions.

48

The genome-wide scores account for only a small fraction of the twin-estimated heritable variance. While the GCTA estimates account for considerably more (21-45%), there remains a majority of the heritable variance to be explained. There is much conjecture about the source of remaining additive genetic variance, including non-additive or rare SNP effects, additive and non-additive structural variation (e.g., CNVs, insertions/deletions), gene-environment interaction, or gene-gene interaction (Manolio, et al., 2009; Zuk, Hechter, Sunyaev, & Lander, 2012). Further research involving larger samples and more comprehensive genotyping, such as whole genome sequencing, will be necessary to tackle these issues.

Future work would also do well to continue to evaluate both individual SNP effects as well as aggregate effects, as both can inform about the genetic architecture of, and genetic relationships among, various psychological traits and other phenotypes. Applying diverse methods, such as twin biometry, GCTA, and genome-wide scoring provides an array of useful information. To be maximally informative, consortia would do well to share among themselves more than GWAS p-values, and to report more than just the genome-wide significant values. Better estimates of regression coefficients, based on tens or hundreds of thousands of individuals, can be used to develop much more powerful genome-wide scores (e.g., as in Allen, et al., 2010), which can then be used in developmentally- and environmentally-informative samples, such as the present sample, to test etiological theories of behavior (Vrieze, et al., 2012). Leveraging consortium meta-analytic results is just now beginning, and have been applied to test developmental

trajectories (Vrieze et al., 2011) and etiological theories of substance use development

(Vrieze, McGue, & Iacono, 2012).

Study 3. The interplay of genes and adolescent development in substance use disorders: Leveraging findings from GWAS meta-analyses to test developmental hypotheses about nicotine consumption.

Main effects of individual common variants with associations to complex diseases have been elusive (Manolio, et al., 2009). It appears that individual samples do not possess sufficient power to detect the vast majority of common single nucleotide polymorphisms (SNPs) relevant to a given complex trait of interest. To overcome this hurdle large consortia have coordinated combined analysis over multiple studies with total sample sizes in the tens to hundreds of thousands. These massive samples have produced promising results for a number of complex traits. Perhaps the best example is with human height, where the GIANT consortium has amassed a total sample size of 183,727 Caucasians (Allen, et al., 2010). Consortia have arisen for other traits as well, such as lipid levels (Willer et al., 2008), body mass index (Speliotes, et al., 2010) and, most important for our purposes, tobacco use (Furberg, et al., 2010; Thorgeirsson et al., 2010).

The need for huge sample sizes is in part due to a combination of small effects and the large number of tests conducted in a GWAS (e.g., see Burton et al., 2009). The problem is compounded when studying developmental or environmental moderators of genetic effects (i.e., GxE or GxD designs). Gene x Environment-Wide Interaction Studies (GEWIS), for example, require at least as many tests as GWAS and much larger sample sizes to ensure adequate power for what is expected to be small GxE effects (e.g.,

sample sizes four times that for GWAS [Manolio, Bailey-Wilson, & Collins, 2006; Thomas, 2010a, 2010b]).   There is presently little reason to believe that GEWIS would be more successful than GWAS have been (Caspi, Hariri, Holmes, Uher, & Moffitt, 2010). Add to this the difficulty of finding existing studies with large sample sizes that possess commensurate phenotypic and environmental measures such that meta-analysis or harmonization is difficult to impossible, and the success of exploratory GEWIS becomes less promising. Unless an a priori GxE hypothesis is extremely strong or the environmental effect is severe and/or uncommon (Caspi, et al., 2010), it is preferable to filter SNPs from genome-wide arrays according to criteria that are likely to result in a subset of SNPs that have promise of environmental interaction (Sullivan et al., 2009). GxE tests can then commence on the presumably much smaller subset of filtered SNPs.

One simple filter is using only SNPs that have shown a main effect (Sullivan, et al., 2009; Thomas, 2010a, 2010b). SNPs with main effects are more likely than other SNPs to be relevant to disease in the first place and thus also more likely to be differentially relevant conditional on environmental experience. Other potentially useful filtering mechanisms include testing for variance differences conditional on genotype (Pare, Cook, Ridker, & Chasman, 2010) as well as selecting SNPs associated with differences between monozygotic twins (Elashoff, Cantor, & Shain, 1991; Magnus, Berg, Borresen, & Nance, 1981). After filtering, one is left with a subset of SNPs probabilistically enriched for GxE interactions. Even better, filtering could be conducted in one sample (or a meta-sample) and the GxE interaction tested in a separate sample, to

avoid capitalization on chance such as when the filtering mechanism and the GxE testing mechanism are non-independent but performed within the same sample.

While there is less literature on gene x development interactions (GxD), they are conceptually similar to GxE interactions and may suffer many of the same pitfalls described above. However, GxD also presents unique challenges. First, as in GxE, there is the need to filter genome-wide arrays for SNPs or other variants that have promise to show a developmental trend (e.g., use SNPs with demonstrated main effect). Second, longitudinal studies are uncommon (relative to cross-sectional studies), expensive, and often span no more than a few assessments over a few years. Obtaining or combining longitudinal sample sizes large enough to detect novel GxD interactions against the background of a million or more common SNPs is probably not achievable at this time. Third and related, using a cross-sectional approach (e.g., two samples that are of different ages) introduces cohort effects that confound any GxD test but, more importantly, they do not allow for the measure of individual change. That is, a cross-sectional approach with two samples of different ages would allow between-group tests but not change at the individual level (Curran & Bauer, 2011). Longitudinal studies are preferred, but they require specialized statistical methods that account for developmental change nested within subject (for a review of methods in longitudinal analysis of behavior, see Curran & Bauer, 2011; Gibbons, Hedeker, & DuToit, 2010)

Development of Nicotine Use and Addiction

Development and environment are key components of nicotine addiction. As to development, adolescence is a critical time for smoking onset. Most smokers begin

smoking in their teenage years and about 90% of smokers express regret about ever starting (Fong et al., 2004). Smoking-related deaths are top causes of morbidity (Center for Disease Control and Prevention, 2005). Those who have not smoked by age 19 are unlikely to become life-long smokers, and have much lower rates of smoking-related morbidity (Curry, Mermelstein, & Sporer, 2009). On the environmental side, remarkable shifts in public policy have contributed to a stark decrease in smoking among adolescents (Monitoring the Future, 2007; Nelson et al., 1995). This change has been attributed to increased knowledge about harmful effects of smoking, government informational campaigns, advertisement bans, pack warnings, cigarette taxes, severe restrictions on where tobacco can be used, and more (Cummings, Fong, & Borland, 2009).

Individuals who never use cigarettes will never be addicted to them. This simple fact results in a host of complications for the longitudinal measurement of developmentally-specific genetic influences. One complication is that initiation, regular use, and addiction are all moderately to highly heritable, and are also genetically related (Maes et al., 2004), indicating that some genetic variants are relevant to both initiation of tobacco use as well as eventual addiction. Indeed, using nicotine is also correlated with alcohol dependence, marijuana dependence, and hard drug dependence (Kendler, Jacobson, et al., 2003; Kendler, Prescott, et al., 2003), and these correlations are highest in adolescence and subside during the transition to adulthood (Vrieze, Hicks, Iacono, & McGue, in press). A theoretical account of this pattern of relationships (Iacono, et al., 2008) argues that behavioral disinhibition (e.g., impulsivity, risk-taking) is partly responsible for the observed correlations among cigarette initiation, use and dependence

and, more broadly, among substance use disorders. Adolescents initiate nicotine, alcohol, and drug use in part because they tend to act impulsively. Those who act more impulsively will be more likely to experiment with cigarettes and, once regular users, more likely to smoke more often, thus consuming more cigarettes per day.

While theoretically important, the disinhibitory hypothesis in no way excludes other mechanisms of addiction. For example, some individuals self-medicate with alcohol in response to emotional disturbance (Sareen, Bolton, Cox, & Clara, 2006). Individual variation exists in the structure and function of neurotransmitter systems integral to substance use and addiction, such as dopamine, serotonin, and GABA. In cigarette smoking, thanks to large-scale consortia, there are verified genetic influences on biological systems that influence the number of cigarettes smoked per day (Thorgeirsson, et al., 2010). The meta-analysis used to guide SNP selection in the present study was conducted on a discovery sample of 31,266 smokers from the ENGAGE consortium with replication in 45,691 smokers from the Glaxo Smith Kline (Ox-GSK) and Tobacco and Genetics (TAG) consortia. While many SNPs approached genome-wide significance, and are reported in supplementary materials, the authors report three genome-wide significant hits that include SNPs in genes *CYP2A6* and *CYP2B6* that encode nicotine-metabolizing enzymes (Ray, Lerman, & Tyndale, 2009) as well as in genes that code for nicotinic acetylcholine receptor subunits (*CHRNB3* and *CHRNA6*).

While these genetic association findings (Furberg, et al., 2010; Thorgeirsson, et al., 2010) are clearly important in understanding the genetic and neurological etiology of nicotine use and addiction, they were obtained in a large heterogeneous sample of adults

from multiple studies. We hypothesize that the relative impact of genetic risk is moderated by development, such that children and adolescents are driven to smoke less by nicotine-specific genetic risk, but more by impulsivity/disinhibitory processes. As individuals who experiment with substances age and mature, we expect nicotine-specific processes to become more and more important in addiction. Indeed, Thorgeirsson, et al., (2010) found evidence of this, if only because they did not find considerable overlap between SNPs associated with cigarettes smoked per day and SNPs associated with smoking initiation. In the present study we test the association of the aggregate effect of SNPs identified in Thorgeirsson, et al., (2010) with CPD at ages 14, 17, 20, and 24, providing a direct assessment of the extent to which the relationship is moderated by age.

Furthermore, the meta-analysis lacked tests of discriminant validity, that is, it did not test whether these nicotine-related hits were also relevant for other substance use disorders, or what the mechanisms of that action may be. In the present study we evaluate the relationship between the nicotine SNPs identified by Thorgeirsson, et al., (2010) and alcohol use. There are multiple etiological pathways that would result in an association with nicotine SNPs and alcohol use; we discuss two. First, it may be that the SNPs are not etiologically specific for smoking, but rather are relevant for more pervasive behavioral systems that impact general risk for substance use like disinhibition or impulsivity. Second, if a genotype is related to nicotine use then this, by definition, indicates that an individual with the risk genotype will experience increased environmental exposure to nicotine. That is, the genotype causes smoking, but it also causes the environment of nicotine intake. Some have theorized that environmental

exposure to drugs and alcohol acts as a "gateway," in which the experience of drug effects cause further use of other, possibly "harder" drugs (e.g., see Kandel & Jessor [2002]) . The present study tests whether these known nicotine SNPs are related also to alcohol use. A positive association would tend to support a gateway-type effect of nicotine exposure on alcohol use, especially if a more generalized risk etiology (e.g., disinhibition) can be ruled out. At first blush the method appears paradoxical in that genetic variants, here SNPs, are actually used as proxies for environmental exposure. This approach is called Mendelian randomization (Smith & Ebrahim, 2003, 2005), and has been successfully applied to address environmental effects on metabolic syndrome (Timpson et al., 2005), drug use (Irons, Iacono, Oetting, & McGue, 2012; Irons, et al., 2007), and other diseases (Schatzkin et al., 2009).

Method

Sample

Participants (N = 3231, 52% female) were drawn from the Minnesota Twin Family Study (MTFS), a community-representative longitudinal study of Minnesota families. The study design has been extensively described elsewhere (Iacono & McGue, 2002), and is an accelerated-longitudinal design composed of two cohorts. The younger cohort (N = 2084, 50% female) was first assessed at age 11 during the years 1988-2005. The older cohort was first assessed at age 17 (N = 1147, 55% female) during the years 1989-1996. Members of the younger, 11-year-old cohort were invited to participate in follow-up assessments at age 14 and 17, and twins from both cohorts were invited to participate in follow-up assessments at age 20 and 24, for a total of five assessments

57

spanning 13 years. Cohorts were combined for all analyses. Participants received

modest honoraria for their assessments and written assent or consent was obtained from

all participants, including the parents of minor children. From its inception, the study has

been continuously approved by the University of Minnesota IRB.

Actual ages of assessment were near the target ages. Pooling across cohorts in the

full sample the mean (SD) ages at assessment were 11.78 (0.43), 14.90 (0.54), 17.84

(0.67), 21.09 (0.81), 24.94 (0.90). Among twins eligible for follow-up assessment, rate

of participation was 93.0% at age 14, 87.3% at age 17, 89.4% at age 20, 91.0% at age 24

and 93.6% at age 29.

Analyses were conducted on two samples: 1) the full sample and 2) a subsample

of participants selected for current smoking. In the subsample we selected, for each age,

only those individuals who smoked at that age. All non-smoker's CPD values were set to

missing, according to the rationale that we do not know how many cigarettes they would

smoke, were they in fact a smoker. We selected the subsample of current smokers

because the SNPs used in this study were identified by meta-analysis for their association

with CPD in current smokers. Since we are developmentally extending the meta-analytic

findings, we hewed to their ascertainment method as closely as possible.[1] A participant

was deemed a smoker if they smoked at least once per month for the 12 months prior to

assessment. This threshold was used to maximize the sample size while only keeping

those individuals who were currently smoking to some extent. Stricter thresholds, such

---

[1] Note that we also conducted all analyses with a subsample of individuals who, at any one assessment
where either 1) a current smoker or 2) had been a smoker in the past. Current smoker was defined just as in
the current smoking sample described in the text. To be a past smoker we required that an individual report
smoking on average one cigarette per day for 12 months. The results from this sample did not change in
any significant way.

as averaging 1 cigarette per day, changed the sample size only by 15-50 people, depending on the age of assessment, and had no impact on the overall findings from this report. A similar procedure was followed to identify current drinkers, with the criterion that they drank at least one alcoholic beverage per month in past 12 months.

Measures

Participants at the age-11 and age-14 assessments were assessed with a computerized questionnaire. The exact question was "During the past 12 months, about how many cigarettes did you smoke in a day when you did smoke?" Participants responded on a 7-point scale. Assessments at later ages were conducted in-person by extensively trained research assistants using a modified version of the expanded Substance Abuse Module (SAM; Robins, et al., 1987). The exact question on the SAM was "During the past 12 months, how much have you smoked (tobacco have you used) on a typical day that you smoked (used tobacco)?" Respondents gave an integer number in response. The computerized questionnaire responses were converted to an integer number for concordance with the SAM. CPD measures were log transformed for all analyses (except descriptive analyses) to mitigate the effect of extreme values on parameter estimates. Frequency of drinking alcohol was also measured with a computerized questionnaire ("During the past 12 months, about how many times did you drink alcohol?") at age 11 and 14 and the SAM ("In the past 12 months, how often on average have you drunk any alcohol (had any alcohol to drink)?") at later ages. Participants responded on a 10-point scale that ranged from "Less than once a year" to "3 or more times a day." There was considerably less skew and a much smaller range of

responses for the measure of drinking frequency, and values thus were not log-transformed.

Genotyping and Imputation

SNPs were genotyped for all members of all families in the Minnesota Center for Twin and Family Research (N = 7438) on an Illumina 660quad array using DNA derived from whole blood for approximately 90% of the sample and from saliva samples for the remainder. For quality control purposes, each 96-well plate included DNA from two members of a single CEPH family (rotated across plates) and one duplicate sample. Markers were excluded if (see Miller, et al., submitted, for additional details): 1) they had been identified as a poorly genotyped marker by Illumina; 2) had more than one mismatch in duplicated QC samples; 3) had a call rate < 99%; 4) had a MAF < 1%; 5) had more than 2 Mendelian inconsistencies across families; 6) significantly deviated from Hardy-Weinberg equilibrium at $p < 1e^{-7}$; 7) was an autosomal marker but associated with sex at $p < 1e^{-7}$; 8) had a significant batch effect at $p < 1e^{-7}$; or 8) there were more than 2 heterozygous X chromosome calls for males or mitochondrial calls for anyone. In total, 32,153 (5.7%) of the 559,982 SNP markers were eliminated by these screens, with the majority (3.6%) being eliminated because of low MAF. Samples were eliminated if: 1) they had > 5000 no-calls; 2) had a low GenCall score; 3) had extreme heterozygosity or homozygosity; or 4) represented a sample mix-up or we could not confirm known genetic relationships. In total, 160 (2.2%) of the total genotyped sample of 7438 failed one or more of these criteria, with the majority (1.7%) failing because of low call rate. Of the 3672 twins initially enrolled in the study, 3365 provided usable genotypes, of which 3231

were Caucasian and used for the present analysis. Only those individuals with genotypes were used in the present study analyses. Those twins with genotypes versus those without did not differ significantly on measures of nicotine use (CPD, days smoked per month, or DSM-IIIR nicotine dependence symptomatology) or alcohol use, with the exception that individuals with genotypes drank slightly more at age 24 (Cohen's $d = .22$, $p = .01$).

Of the 92 SNPs used in the SNP score, 28 were available on the Illumina 660W quad array used in the full sample. The remaining SNPs were imputed with MaCH (Li, Willer, Sanna, & Abecasis, 2009; Li, Willer, Ding, Scheet, & Abecasis, 2010) based on the August 2010 reference haplotypes, each with satisfactory imputation quality (mean r^2 = .92; median = .96; SD = .11; range = [0.52, 1.0]).

SNPs used in the molecular genetic analysis were selected from supplemental materials reported in (Thorgeirsson, et al., 2010). Only SNPs in regions of interest were reported; that is, regions that harbored promising SNPS, including all genome-wide significant SNPs, based on the full meta-sample were reported. All SNPs with associations reported for cigarettes smoked per day were used in the present analysis. The SNP set was winnowed on the basis of linkage disequilibrium in an iterative fashion. The most significant SNP was selected, and all SNPs within $r^2 > .7$ of that SNP were removed from the SNP set. Then the next most significant SNP in the reduced dataset was evaluated, and SNPs with $r^2 > .7$ were removed, and so on. We chose a liberal r^2 because our initial SNP set was rather small (602), and we wanted to exclude only the most highly-redundant SNPs. This process retained 92 of the 602 SNPs originally

reported in the supplemental materials. For all analyses, these 92 SNPs were summed according to their meta-analytic regression weights to form a genetic risk score for tobacco use. Information on the final 92 SNPs are available upon request.

Longitudinal Heritability Analysis
The longitudinal model employed was saturated, in that all across-time CPD covariances were simultaneously estimated. Within-family clustering due to twin status was accounted for with standard twin modeling (Neale & Cardon, 1992). This technique decomposes the observed CPD phenotypic variance-covariance matrix into additive genetic variance (A), shared environmental variance (C), and unshared environmental variance (E). Because so little variance is observed for CPD at age 11, our longitudinal models only included measurements at age 14, 17, 20, and 24. All heritability models were fit with OpenMx 1.0.6 (Boker, et al., 2011) of the *R* Environment version 2.13.1 (R Development Core Team, 2011). Confidence intervals were computed using non-linear constraints and likelihood ratio tests according to standard practice (Miller & Neale, 1995; Neale & Miller, 1997). All models were fit under full information maximum likelihood to account for missing data. Covariates used in these analyses were year of birth, sex, and cohort status. Covariates were incorporated as fixed effects as described in the next section.

Incorporating the SNP Score in the Longitudinal Design
The genetic SNP score can be incorporated directly into the ACE model described above as a fixed effect. This approach is equivalent to multivariate linear regression except that the model accounts for clustering within families (twin pairs are nested within

families) and within subjects over time due to the repeated measurements. Resulting regression coefficient estimates and standard errors are corrected for the non-independent observations. The overall model is a special case of a mixed effects regression (Pinheiro & Bates, 2000), where the variance components are random effects and the SNP score and covariates are fixed effects, with the difference being that we used a variance-covariance matrix structured to account for twin zygosity.

Year of birth, cohort, sex, age of assessment, and the first ten genetic principal components were used as covariates in all SNP score association tests. Genetic principal components account for ancestry effects and were based on a subsample of 10,000 SNPs from sample founders (i.e., unrelated subjects) and computed in Eigenstrat (Price, et al., 2006). Scripts are available upon request.

Mendelian Randomization

Mendelian Randomization is a technique by which one can use genetic variation as a proxy for environmental exposure, and has been the subject of several reviews (Ebrahim & Smith, 2008; Smith, 2011). In the present study we use genetic variation associated with nicotine use as a proxy for the environmental exposure to nicotine. That is, individuals with high-risk genetic variation will be more likely to smoke more cigarettes. If the extent of nicotine use causes later alcohol use (a plausible prediction to be made from the gateway model), then we expect these nicotine-relevant variants to also be associated with alcohol use, even though the strength of the relationship may be attenuated to the extent expected from the less than perfect correlation between alcohol

and nicotine use. All alcohol analyses were conducted in the same way as described for nicotine above.

Results

Descriptive Statistics

Descriptive statistics are provided in Table 6 for both the full sample and the subsample of current smokers, including sample sizes, number of twin pairs, and mean (SD) of CPD. In the full sample there are around 1000 or more participants assessed at each of the ages. In the subsample the number of participants is reduced considerably, especially at younger ages. At older ages the samples of current smokers are much larger, comprising 1/2 to 3/4 of the full sample. Our criteria for current smoking was intended to be inclusive, and to capture all individuals who were smoking cigarettes even to only a small extent (e.g., smoking on average one cigarette per month) during the assessment period. Individual trends are reported graphically in Figure 7, along with a running mean and standard deviation (based on overlapping windows of 500 subjects). Very few 11-year-olds smoke to any extent. Some individuals dramatically increase their intake at age 14, but again the vast majority uses no cigarettes (the median at all young ages is zero). By the age 17 assessment the mean CPD increases to around 2 for females and 3 for males. By age 20 and 24 the mean is around 4 for females and 6 for males. As observed in the figure, the standard deviation of CPD increases with the mean, a consequence of the fact that most individuals use zero to one CPD regardless of age. The "dip" at around age 17, visible for males and females, appears to represent the switch from computerized assessment of CPD to in-person interview assessment. Another

possibility is that the dip is due to a cohort effect, as the older cohort begins their

assessment at age 17, and may have been using cigarettes to a less extent than the

younger cohort, resulting in the observed dip. Plotting a running mean only on the

younger cohort, excluding the older cohort, yields the same age-17 dip in CPD, and

suggests that the decline in CPD is due to the change in assessment format.

Table 6. Descriptive Statistics.

| Sample	Measure	Sex	Age of Assessment 11	14	17	20	24
Full Sample	N	M	1041	991	1269	999	1053
		F	1043	999	1485	1220	1121
	Full MZ pairs	M	337	319	409	319	329
		F	325	310	465	379	339
	Full DZ pairs	M	170	160	194	146	161
		F	184	175	250	209	183
	CPD Mean (SD)	M	0.14 (.50)	1.03 (2.75)	3.41 (6.31)	6.12 (8.03)	6.13 (8.18)
		F	0.07 (.39)	0.80 (2.74)	2.22 (4.87)	3.68 (6.05)	3.19 (5.61)
	Drinking Frequency Mean (SD)	M	0.08 (0.34)	0.69 (1.27)	2.14 (1.77)	4.05 (1.66)	4.37 (1.59)
		F	.03 (.20)	0.65 (1.23)	1.84 (1.70)	3.28 (1.61)	3.67 (1.53)
Subsample of Current Smokers	N	M	125	367	709	766	793
		F	57	283	625	681	545
	Full MZ pairs	M	27	100	188	213	204
		F	6	67	137	143	108
	Full DZ pairs	M	7	32	77	101	111
		F	10	34	84	96	64
	CPD Mean (SD)	M	1.16 (0.95)	2.79 (3.94)	6.10 (7.40)	7.98 (8.32)	8.13 (8.51)
		F	1.21 (1.21)	2.81 (4.56)	5.27 (6.35)	6.60 (6.80)	6.57 (6.53)
Subsample of Current Drinkers	N	M	12	184	798	916	988
		F	3	199	840	1075	1025
	Full MZ pairs	M	0	37	212	277	299
		F	0	34	207	309	289
	Full DZ pairs	M	0	8	88	129	145
		F	0	22	118	170	158
	Drinking Frequency Mean (SD)	M	2.42 (0.67)	3.01 (1.20)	3.25 (1.26)	4.38 (1.29)	4.62 (1.27)
		F	2.33 (0.58)	2.82 (1.13)	3.06 (1.23)	3.67 (1.27)	3.95 (1.28)

CPD = Cigarettes Smoked per Day
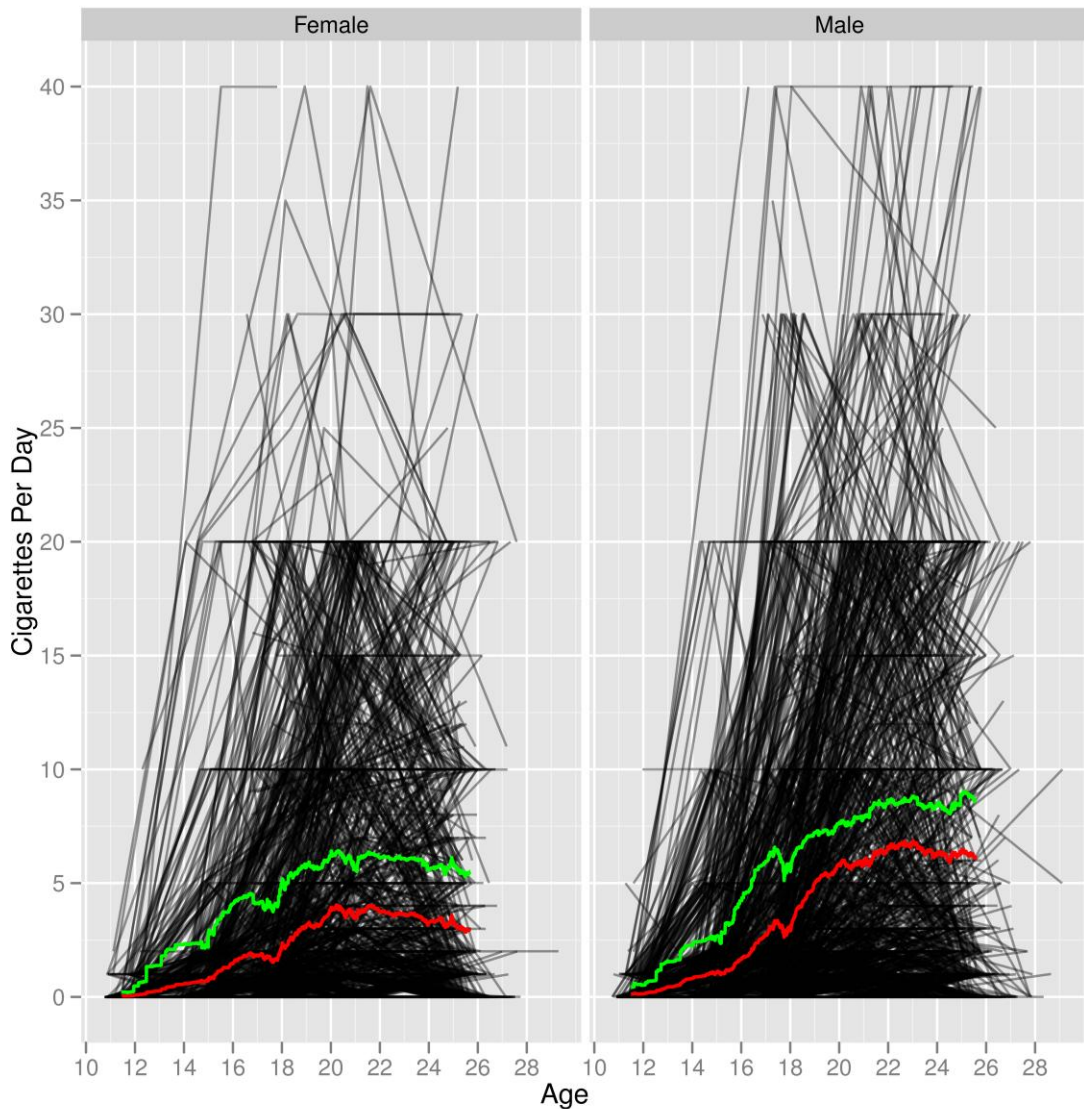
Figure 7. Developmental Trajectory of Cigarettes Smoked per Day.
The red line is a running mean. The green line is a running standard deviation. The dip at approximately age 17 is when the assessment protocol switched from a computerized questionnaire to an in-person interview.

Twin Heritability Model Results

   Variance components from the biometric twin model are reported in Figure 8 as a

series of correlation matrices with superimposed heatmaps. The phenotypic correlation

matrix is denoted "P." As the participants age, CPD becomes more stable over time. The

correlation between age 14 and age 17 CPD is .55; the correlation between age 20 and

67

age 24 CPD is .78. Note that the A, C, and E matrices are scaled such that the A, C, and E matrix, if summed element-wise, give the P matrix. The "A" matrix denotes that portion of the phenotypic matrix due to the additive effect of genes. It follows the same pattern as the phenotypic matrix, with genetic effects becoming more stable over time. Shared environmental effects are stronger at younger ages and generally taper over time. Unshared environmental effects contribute largely only to variance at assessments, but also contributes increasingly to stability of CPD across assessments, perhaps as a result of the highly addictive nature of cigarette smoking.
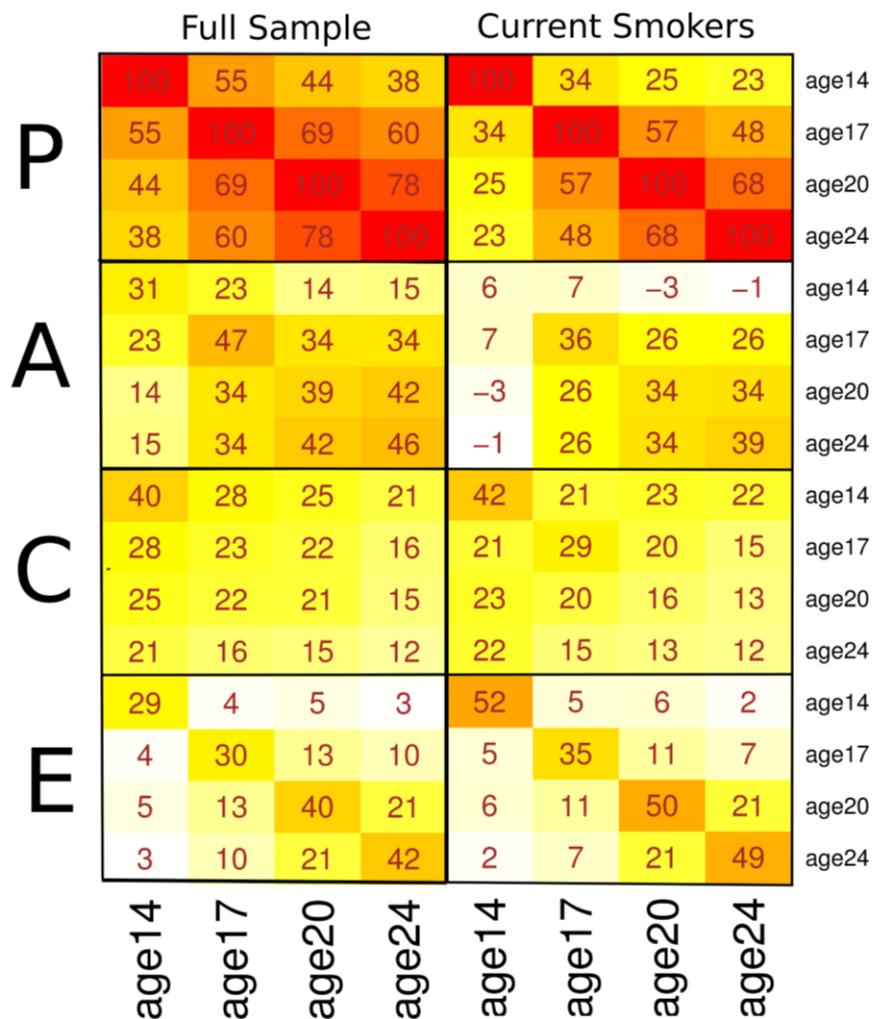
	Full Sample				Current Smokers				
P	100	55	44	38	100	34	25	23	age14
	55	100	69	60	34	100	57	48	age17
	44	69	100	78	25	57	100	68	age20
	38	60	78	100	23	48	68	100	age24
A	31	23	14	15	6	7	−3	−1	age14
	23	47	34	34	7	36	26	26	age17
	14	34	39	42	−3	26	34	34	age20
	15	34	42	46	−1	26	34	39	age24
C	40	28	25	21	42	21	23	22	age14
	28	23	22	16	21	29	20	15	age17
	25	22	21	15	23	20	16	13	age20
	21	16	15	12	22	15	13	12	age24
E	29	4	5	3	52	5	6	2	age14
	4	30	13	10	5	35	11	7	age17
	5	13	40	21	6	11	50	21	age20
	3	10	21	42	2	7	21	49	age24
	age14	age17	age20	age24	age14	age17	age20	age24	

Figure 8. Heatmap of Correlation Matrix and Heritability Components of CPD for the Full Sample and Subsample of Current Smokers.

The figure is composed of two columns—the full sample is on the left and the subsample of current smokers is on the right. The upper two 4x4 matrices denoted "P" are the total phenotypic correlation matrices for each sample. Below that are the additive genetic (A), shared environmental (C), and unshared environmental (E) components for each sample. The matrices are scaled such that element-wise summation of the A, C, and E matrices gives the phenotypic (P) correlation matrix. This scaling gives the CPD heritabilities and environmental effects along the diagonal of the A, C, and E matrices (e.g., CPD is 31% heritable at age 14 in the full sample).

Including the SNP score in the Model

The SNP score was tested in both the full sample and the subset of current smokers. The score effect was non-significant at each age of assessment for CPD in the full sample. Results are not shown here for lack of space.

The SNP score did, however, show significant effects in the subsample of current smokers, consistent with the ascertainment method used by the meta-analysis from which the SNP score was derived. Results are given in Table 7. The effect was strongest for later ages with regression coefficients of .021 and .014 at ages 20 and 24, accounting for 1.0% of the variance at age 20 and 0.4% at age 24. While the score was not significant for younger ages, the regression coefficients were in the same direction but of a smaller magnitude (.010 and .009 at age 14 and 17). The SNP score at age 14 and 17 accounted for 0.1% to 0.4% of the variance, respectively, suggesting an attenuation of effect instead of the lack of effect. A composite test of the null hypothesis that the regression coefficient was the same at all ages approached statistical significance ($\chi^2 = 6.2$, df =3, $p = .10$).

Table 7. SNP Score Effects on CPD and Alcohol Use Frequency by Age.

Drug	Age	N	Coefficient	95% C.I.	LRT	p-value	Variance Accounted for by Score
Nicotine	14	650	.010	(-.002, .023)	2.64	.10	0.38%
	17	1334	.009	(-.003, .022)	2.28	.13	0.08%
	20	1447	.021	(.009, .034)	10.86	.001	0.97%
	24	1338	.014	(.001, .023)	4.15	.04	0.38%
Alcohol	14	383	-.003	(-.010, .004)	0.60	.44	0.16%
	17	1638	.002	(-.002, .006)	0.84	.36	0.10%
	20	1991	-.001	(-.004, .003)	0.27	.61	0.03%
	24	2013	-.004	(-.007, -.001)	5.57	.02	0.37%

LRT is the likelihood ratio test statistic, distributed as χ^2 on 1 degree of freedom, of the model with the score effect coefficient estimated, versus the model where it is fixed to zero. Measures of nicotine use were log-transformed for this analysis. Coefficients are unstandardized regression weights, with negative values indicating decreased risk for substance use.

In addition to accounting for significant variance at age 20 and 24, the SNP score also accounted for significant covariance (1.0%, $p < .05$, as tested by 200 bootstrap replications) between CPD at age 20 and 24, indicating that the SNP score is not only associated with CPD at 20 and 24, but also associated with stability in CPD for those individuals smoking between ages 20 and 24.

Also listed in Table 7 is the SNP score relationship with frequency of alcohol use in the subsample of current alcohol users. As is clear from the table, the SNP score was unrelated to frequency of alcohol use at ages 14, 17, and 20. It was, however, significantly associated at age 24, but in the opposite direction as that observed for nicotine use. That is, at age 24 higher SNP scores were associated with greater CPD in smokers but less frequent drinking in drinkers. The SNP score also showed no

association with alcohol use in the full sample or subsample of smokers (results not shown).

Discussion

This report is among the first to place SNP main effects identified by meta-analysis in a longitudinal developmental context (see also Vrieze, et al., 2011). This approach can refine the nature of the meta-analytic association, and allow novel tests of existing addiction theory. We obtained three important findings. First, CPD is moderately heritable (30% to 50%) in the full sample and the subsample of current smokers. Genes account for a majority of the stability of smoking across time, especially at later ages. Shared environment accounts for the majority of stability from age 14 to age 17, indicating the potential for significant gene and environmental main effects moderated by development in CPD.

Second, the SNP score was not statistically significantly associated with cigarettes smoked per day at age 14 or 17, but was significantly associated at age 20 and 24. If replicable, this indicates that these genetic variants are less relevant during adolescence, when a large number of individuals begin experimenting with cigarettes and/or regularly using cigarettes. Differences in effect sizes at the different ages approached statistical significance, but requires replication. This suggests the possibility of a GxD effect, and that other mechanisms, such as behavioral disinhibition, play significantly larger roles during adolescence with CPD, consistent with literature cited in the introduction. In addition to significant associations with CPD at age 20 and 24, the longitudinal study allowed us to test whether the score accounted for stability in CPD across age. The score

accounted for about 1% of the covariance between age 20 and age 24 smoking, and is thus relevant to the maintenance of CPD over this four year interval.

Third, the SNP score was not significantly related to alcohol use at ages 14, 17, and 20, indicating that these SNPs, even though they increase the likelihood of smoking at ages 20 and 24, play no detectable role in alcohol use. The sensitivity of this analysis is limited by the relatively small genetic effect on nicotine use ($r^2 < 1\%$) as well as the small sample size, and requires confirmation in larger samples. While the present results do not rule out the existence of a positive effect of environmental nicotine exposure on alcohol use, such a hypothesized effect is too small to be detected in this sample. There was a significant SNP score effect on alcohol use at age 24, but in the opposite direction as that observed for cigarettes. This result was unexpected and while possibly spurious, nevertheless fails to support the notion that the nicotine SNP score increase risk for alcohol use. One might conjecture that it represents a trend for early adults to specialize in their drug use habits and use available resources (e.g., money, time) for drugs of choice at the expense of other substances. However, in an analysis of the SNP score's relationship with alcohol use frequency in the subsample of current smokers, there was no significant effect at age 24 (regression coefficient = -0.018, $\chi^2 = 1.3$, df = 1, p = 0.25), suggesting that the effect in general may well be spurious.

The results of this study are preliminary, and require replication and extension, but one can begin to see the value of evaluating genetic effects on substance use in a developmental and environmental context. As sample sizes grow, and investigators begin sharing data at higher rates, tests of developmental and environmental moderation

become more powerful and feasible. Mendelian randomization tests offer an elegant quasi-experimental way to test the causal role of an environmental exposure, and will become a mainstay in quasi-experimental research.

The novelty of this study requires a number of cautionary remarks. As noted in the introduction, GxD is conceptually similar to GxE, and likely shares many of the same pitfalls. One major hurdle is filtering candidate SNPs that show potential for GxD, and this is where the current study derives much of its strength. The longitudinal sample used in this report is quite small by GWAS standards, and could not by itself be used profitably to discover reliable GxD interactions against a background of 1 million SNPs. By leveraging results from the meta-analytic literature, which used tens of thousands of subjects, we were able to focus our GxD design only on highly promising SNPs that have high potential for replication.

Even by leveraging meta-analytically identified SNPs, the present sample was not large enough to identify developmental changes in individual SNP effects but, rather, SNPs had to be combined into a single genetic risk score before the effect was large enough to be detectable and testable in this sample. This score appears to be biologically coherent, in that the effects tagged by the SNPs are relevant to either nicotine metabolism or nicotine transmission. Future work may do better by combining scores within explicit biological pathways through the use of pathway databases (Vink et al., 2009). In this way, genetic scores take on more refined biological meaning and results provide clearer theoretical import. A major challenge to the pathway approach is simply in identifying enough SNPs or other genetic variants within a pathway that have measureable effects on

the phenotype. While non-trivial and etiologically informative, it is very unlikely that this SNP score could be utilized in any way in a personalized medicine context. There are many theories as to why no sizable genetic effects have been discovered for complex traits (Manolio, et al., 2009). Among these are the possibility that complex traits arise from polygenic inheritance where large numbers of common SNPs have individually small effects but account for considerable variance in aggregate (e.g., see Allen, et al., 2010; Visscher, Yang, & Goddard, 2010; Yang, et al., 2010). Other possibilities include rare variants that may have individually large effects, such as rare nonsense or missense SNPs or copy-number variants such as insertions/deletions, or variants related to gene expression and regulation (e.g., see Altshuler et al., 2010; Montgomery, Lappalainen, Gutierrez-Arcelus, & Dermitzakis, 2011). Undoubtedly, future work will be able to combine common tag SNPs from genome-wide arrays with rare and structural variation from sequencing into much more statistically powerful and biologically interpretable risk scores.

A limitation of our approach is that it relies on consortia to produce candidate SNPs for developmental analysis. This ignores the possibility of developmentally-relevant SNPs that result in similar adult status. To use a concrete example, there may exist powerful genetic variants that affect the onset of the pubertal growth spurt, and thus affect the shape of an individual's growth trajectory, but have little or no influence on their eventual adult height. This SNP would never be identified by meta-analyses of adult height, but is extremely important in understanding how growth culminates in adult height. Just as in the GWAS literature, consortia of longitudinal studies are required, and

75

individual investigators must share their data in order to achieve large enough sample sizes for more meaningful comparisons.

Substance use development is confounded with environmental change during adolescence. As youths age it is arguable that their home environment changes to accommodate an increased expectation of autonomy, and increased needs for independence. In the United States, for example, one can legally drive an automobile at age 16, legally purchase tobacco products at 18, and legally purchase alcohol at 21, all of which are within the time interval under investigation in this study. While a caregiver may restrict automobile or substance use, it is clear that developmental maturation can be confounded with dramatic environmental shifts in substance availability and social norms surrounding use. In fact, this observation can easily explain the current findings. Adolescents at age 14 or 17 may have a genotype that increases their CPD but are unable to satisfy that risk because cigarettes are not readily available to them, thus the SNP score's effect is moderated not by developmental neurological or biological maturation, but by a developmentally-confounded environment (i.e., development serves as a proxy for environmental change). Add to this the extra complication that individuals with riskier genotypes are more likely to have biological parents with riskier genotypes, who will smoke more than those who do not (i.e., gene-environment correlation [Jaffee & Price, 2007]), and one can appreciate the challenges we face in understanding how genes and environment affect the development of addiction.

The present study evaluated the developmental moderation of cigarettes smoked per day by a SNP score derived from a meta-analysis of current smokers. The results

suggest developmental moderation with an increased effect at age 20 and 24, and decreased effect at age 14 and 17. These results are consistent with the notion that nicotine-specific substance use risk is less important at younger ages, and becomes more important as individuals age into adulthood. The SNP score was unrelated to alcohol use at any age, indicating that the SNP score effects are specific to nicotine, and do not tap more general etiological processes simultaneously relevant to nicotine and alcohol use

Summary

Behavioral disinhibition is a trait theorized to affect one's propensity to experiment with substances, as well as make it harder for one to cut down or stop their use. We hypothesized throughout that adolescents are especially vulnerable to substance use initiation and use, in part because they are disinhibited relative to adults. What is more, twin and adoption studies have shown that behavioral disinhibition and substance use measures are largely heritable, as are their inter-correlations. The first study found that correlations among substance use disorders decline with age from 14 to 29, with a majority of the drop occurring between age 17 and 24. If disinhibition is highest during adolescence, and disinhibition plays a causal role in substance use, then we expect substance use correlations to be highest at that age. As individuals mature and become more constrained in their behavior, then we expect indiscriminant use to resolve. We found precisely these trends. What is more, the genetic contribution to indiscriminant substance use diminished with age while the non-shared contribution increased, suggesting that individual addictive processes caused by nonshared environment increasingly contribute to substance use as individuals age into their late 20's.

77

The second study was an attempt to go beyond twin- and adoption-based studies of the genetic etiology and genetic links among behavioral disinhibition and substance use. GWAS in a sample of this size (effective N < 7,000) is almost certain to fail to identify relevant genes (Visscher, et al., 2012). To overcome sample size limitations, we attempted a genome-wide scoring approach, which adds up SNP effects across the entire genome. We then compared the scoring results it to biometric heritability estimates as well as SNP-based heritabilities estimated by GCTA. We found that the biometric and GCTA estimates were highly concordant and mutually informative, and that genome-wide scoring identified measureable traces of genetic association in this sample. While limited, we safely conclude that there were small but detectable pleiotropic SNP associations with various measures of substance use (nicotine, alcohol, and drug), as well as between a measure of behavioral disinhibition and substance use.

The last study was another attempt to use measured genetic variants (SNPs) to understand genetic and, through a Mendelian randomization design, environmental etiology on nicotine and alcohol use. We used SNPs known to be associated with cigarettes smoked per day (CPD), and evaluated how their association with CPD changed during adolescence. The genetic effects were smaller during adolescence (age 14 and 17) and grew in strength over time. The results suggest that adolescent environments protect against risky genotypes. The Mendelian randomization study (Ebrahim & Smith, 2008) tested whether individuals genetically prone to smoke, and therefore more likely on average to have the environmental experience of cigarette smoking, were more likely to use more alcohol. No such association was found, in that those more likely to smoke

were no more likely to use alcohol to a greater extent, contrary to what a gateway theory of addiction would predict. The results are limited by the extent of the association between the SNPs and smoking, which in this case was small ($r^2 < 1\%$). As more risk variants are discovered, the utility of the Mendelian randomization design will increase.

We are optimistic of the future of genetic research on substance use. Individual investigators are learning about the necessity of data sharing and cross-study cooperation. As larger samples are collected more genes will be discovered, which will facilitate the discovery of yet more genes. Individual investigators with refined study designs will then be able to study in more detail the actual etiological mechanisms by which genes interact with environment and development to produce the clinical and behavioral phenotypes with which we are all so familiar. The ultimate goal, of course, is to understand etiology in order to answer longstanding nosological questions about mental illness and inform effective scientifically sound and effective intervention.

Bibliography

Allen, H. L., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F., et al.
(2010). Hundreds of variants clustered in genomic loci and biological pathways
affect human height. *Nature, 467*(7317), 832-838.

Altshuler, D. L., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark,
A. G., et al. (2010). A map of human genome variation from population-scale
sequencing. *Nature, 467*(7319), 1061-1073.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental
disorders (3rd ed., rev.)*. Washington, D.C.: Author.

Aulchenko, Y. S., Ripke, S., Isaacs, A., & van Duijn, C. M. (2007). GenABEL: an R
library for genome-wide association analysis. *Bioinformatics, 23*(10), 1294-1296.

Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist
drugs: a neurocognitive perspective. *Nature Neuroscience, 8*(11), 1458-1463.

Bentler, P. M. (1990). Comparative Fit Indexes in Structural Models. *Psychological
Bulletin, 107*(2), 238-246.

Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., et al. (2011). OpenMx:
An Open Source Extended Structural Equation Modeling Framework.
Psychometrika, 76(2), 306-317.

Breiman, L., & Spector, P. (1992). Submodel Selection and Evaluation in Regression -
the X-Random Case. *International Statistical Review, 60*(3), 291-319.

Brown, T. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: The
Guilford Press.

Burton, P. R., Hansell, A. L., Fortier, I., Manolio, T. A., Khoury, M. J., Little, J., et al. (2009). Size matters: just how big is BIG? Quantifying realistic sample size requirements for human genome epidemiology. *International Journal of Epidemiology, 38*(1), 263-273.

Caruso, J. C., & Cliff, N. (1997). Empirical size, coverage, and power of confidence intervals for Spearman's rho. *Educational and Psychological Measurement, 57*(4), 637-654.

Casey, B. J., Getz, S., & Galvan, A. (2008). The adolescent brain. *Developmental Review, 28*(1), 62-77.

Casey, B. J., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends in Cognitive Sciences, 9*(3), 104-110.

Caspi, A., Hariri, A. R., Holmes, A., Uher, R., & Moffitt, T. E. (2010). Genetic Sensitivity to the Environment: The Case of the Serotonin Transporter Gene and Its Implications for Studying Complex Diseases and Traits. *American Journal of Psychiatry, 167*(5), 509-527.

Center for Disease Control and Prevention. (2005). Annual Smoking-attributable mortality, years of potential life lost, and productivity losses. *Morbidity and Mortality Weekly Report, 54*(25), 3.

Cummings, K. M., Fong, G. T., & Borland, R. (2009). Environmental Influences on Tobacco Use: Evidence from Societal and Community Influences on Tobacco Use and Dependence. *Annual Review of Clinical Psychology, 5*, 433-458.

Curran, P. J., & Bauer, D. J. (2011). The Disaggregation of Within-Person and Between-Person Effects in Longitudinal Models of Change. *Annual Review of Psychology, Vol 62, 62*, 583-619.

Curry, S. J., Mermelstein, R. J., & Sporer, A. K. (2009). Therapy for Specific Problems: Youth Tobacco Cessation. *Annual Review of Psychology, 60*, 229-255.

Ebrahim, S., & Smith, G. D. (2008). Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology? *Human Genetics, 123*(1), 15-33.

Elashoff, J. D., Cantor, R. M., & Shain, S. (1991). Power and Validity of Methods to Identify Variability Genes. *Genetic Epidemiology, 8*(6), 381-388.

Fong, G. T., Hammond, D., Laux, F. L., Zanna, M. P., Cummings, K. M., Borland, R., et al. (2004). The near-universal experience of regret among smokers in four countries: Findings from the International Tobacco Control Policy Evaluation Survey. *Nicotine & Tobacco Research, 6*, 341-351.

Franke, A., McGovern, D. P. B., Barrett, J. C., Wang, K., Radford-Smith, G. L., Ahmad, T., et al. (2010). Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nature Genetics, 42*(12), 1118-+.

Furberg, H., Kim, Y., Dackor, J., Boerwinkle, E., Franceschini, N., Ardissino, D., et al. (2010). Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nature Genetics, 42*(5), 441-134.

Gibbons, R. D., Hedeker, D., & DuToit, S. (2010). Advances in Analysis of Longitudinal Data. *Annual Review of Clinical Psychology, Vol 6, 6*, 79-107.

Goldstein, R. Z., & Volkow, N. D. (2002). Drug addiction and its underlying

    neurobiological basis: Neuroimaging evidence for the involvement of the frontal

    cortex. *American Journal of Psychiatry, 159*(10), 1642-1652.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning :*

    data mining, inference, and prediction (2nd ed.). New York, NY: Springer.

Hicks, B. M., Schalet, B. D., Malone, S. M., Iacono, W. G., & McGue, M. (2011).

    Psychometric and genetic architecture of substance use disorder and behavioral

    disinhibition measures for gene association studies. *Behavior Genetics, 41*(4),

    459-475.

Hirschhorn, J. N., & Daly, M. J. (2005). Genome-wide association studies for common

    diseases and complex traits. *Nature Reviews Genetics, 6*(2), 95-108.

Hu, L. T., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance

    Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural*

    Equation Modeling-a Multidisciplinary Journal, 6(1), 1-55.

Iacono, W. G., Carlson, S. R., Taylor, J., Elkins, I. J., & McGue, M. (1999). Behavioral

    disinhibition and the development of substance use disorders: Findings from the

    Minnesota Twin Family Study. *Development and Psychopathology, 11*, 869-900.

Iacono, W. G., Malone, S. M., & McGue, M. (2008). Behavioral disinhibition and the

    development of early-onset addiction: Common and specific influences. *Annual*

    Review of Clinical Psychology, 4, 325-348.

Iacono, W. G., & McGue, M. (2002). Minnesota Twin Family Study. *Twin Research,*

    5(5), 482-487.

Irons, D. E., Iacono, W. G., Oetting, W. S., & McGue, M. (in press). Developmental
trajectory and environmental moderation of the effect of ALDH2 polymporphism
on alcohol use. *Alcoholism: Clinical and Experimental Research.*

Irons, D. E., McGue, M., Iacono, W. G., & Oetting, W. S. (2007). Mendelian
randomization: A novel test of the gateway hypothesis and models of gene-
environment interplay. *Development and Psychopathology, 19*(4), 1181-1195.

Jaffee, S. R., & Price, T. S. (2007). Gene-environment correlations: A review of the
evidence and implications for prevention of mental illness. *Molecular Psychiatry,
1*(1), 1-11.

Kandel, D. B., & Jessor, R. (2002). The gateway hypothesis revisited. In D. B. Kandel
(Ed.), *Stages and Pathways of Drug Involvement*. New York: Cambridge
University Press.

Kendler, K. S., Jacobson, K. C., Prescott, C. A., & Neale, M. C. (2003). Specificity of
genetic and environmental risk factors for use and abuse/dependence of cannabis,
cocaine, hallucinogens, sedatives, stimulants, and opiates in male twins. *American
Journal of Psychiatry, 160*(4), 687-695.

Kendler, K. S., Prescott, C. A., Myers, J., & Neale, M. C. (2003). The structure of genetic
and environmental risk factors for common psychiatric and substance use
disorders in men and women. *Archives of General Psychiatry, 60*(9), 929-937.

Kessler, R. C., Chiu, W. T., Demler, O., Merikangas, K. R., & Walters, E. E. (2005).
Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the

national comorbidity survey replication. (vol 62, pg 617, 2005). *Archives of General Psychiatry, 62*(7), 709-709.

Klimstra, T. A., Hale, W. W., Raaijmakers, A. W., Branje, S. J. T., & Meeus, W. H. J. (2009). Maturation of Personality in Adolescence. *Journal of Personality and Social Psychology, 96*(4), 898-912.

Krueger, R. F., Hicks, B. M., Patrick, C. J., Carlson, S. R., Iacono, W. G., & McGue, M. (2002). Etiologic connections among substance dependence, antisocial behavior, and personality: Modeling the externalizing spectrum. *Journal of Abnormal Psychology, 111*(3), 411-424.

Krueger, R. F., & Markon, K. E. (2006). Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology, 2*, 111-133.

Leckman, J. F., Scholomskas, D., Thompson, W. D., Belanger, A., & Weisman, M. M. (1982). Best estimate of lifetime psychiatric diagnosis: A methdlogical study. *Archives of General Psychiatry, 39*, 879-883.

Lenroot, R. K., Gogtay, N., Greenstein, D. K., Wells, E. M., Wallace, G. L., Clasen, L. S., et al. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *Neuroimage, 36*(4), 1065-1073.

Li, Y., Willer, C., Sanna, S., & Abecasis, G. (2009). Genotype Imputation. *Annual Review of Genomics and Human Genetics, 10*, 387-406.

Li, Y., Willer, C. J., Ding, J., Scheet, P., & Abecasis, G. R. (2010). MaCH: Using Sequence and Genotype Data to Estimate Haplotypes and Unobserved Genotypes. *Genetic Epidemiology, 34*(8), 816-834.

Luczak, S. E., Glatt, S. J., & Wall, T. L. (2006). Meta-analyses of ALDH2 and ADH1B with alcohol dependence in Asians. *Psychological Bulletin, 132*(4), 607-621.

Maes, H. H., Sullivan, P. F., Bulik, C. M., Neale, M. C., Prescott, C. A., Eaves, L. J., et al. (2004). A twin study of genetic and environmental influences on tobacco initiation, regular tobacco use and nicotine dependence. *Psychological Medicine, 34*(7), 1251-1261.

Magnus, P., Berg, K., Borresen, A. L., & Nance, W. E. (1981). Apparent Influence of Marker Genotypes on Variation in Serum-Cholesterol in Monozygotic Twins. *Clinical Genetics, 19*(1), 67-70.

Manolio, T. A., Bailey-Wilson, J. E., & Collins, F. S. (2006). Opinion - Genes, environment and the value of prospective cohort studies. *Nature Reviews Genetics, 7*(10), 812-820.

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., et al. (2009). Finding the missing heritability of complex diseases. *Nature, 461*(7265), 747-753.

Meehl, P. E. (1992). Factors and Taxa, Traits and Types, Differences of Degree and Differences in Kind. *Journal of Personality, 60*(1), 117-174.

Meehl, P. E. (2001). Comorbidity and taxometrics. *Clinical Psychology-Science and Practice, 8*(4), 507-519.

Miller, M. B., Basu, S., Cunningham, J., Oetting, W., Schork, N. J., Iacono, W. G., et al. (submitted). The Minnesota Center for Twin and Family Research Genome-Wide Association Study.

Miller, M. B., & Neale, M. C. (1995). Using Nonlinear Constraints in Mx to Automate Estimation of Confidence-Intervals for Parameters in Genetic Models. *Behavior Genetics, 25*(3), 279-280.

Moffitt, T. E., Caspi, A., Taylor, A., Kokaua, J., Milne, B. J., Polanczyk, G., et al. (2010). How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. *Psychological Medicine, 40*(6), 899-909.

Monitoring the Future. (2007). Trends in cigarette smoking and smokeless tobacco: University of Michigan.

Montgomery, S. B., Lappalainen, T., Gutierrez-Arcelus, M., & Dermitzakis, E. T. (2011). Rare and common regulatory variation in population-scale sequenced human genomes. *PLoS Genetics, 7*(7), e1002144.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dortrecht, The Netherlands: Kluwer Academic.

Neale, M. C., & Miller, M. B. (1997). The use of likelihood-based confidence intervals in genetic models. *Behavior Genetics, 27*, 113-120.

Nelson, D. E., Giovino, G. A., Shopland, D. R., Mowery, P. D., Mills, S. L., & Eriksen, M. P. (1995). Trends in Cigarette-Smoking among Us Adolescents, 1974 through 1991. *American Journal of Public Health, 85*(1), 34-40.

Pare, G., Cook, N. R., Ridker, P. M., & Chasman, D. I. (2010). On the Use of Variance per Genotype as a Tool to Identify Quantitative Trait Interaction Effects: A Report from the Women's Genome Health Study. *PLoS Genetics, 6*(6).

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics, 38*(8), 904-909.

R Development Core Team. (2011). R: A language and environment for statistical computing.

Ray, R., Lerman, C., & Tyndale, R. F. (2009). Nicotine Dependence Pharmacogenetics: Role of Genetic Variation in Nicotine-Metabolizing Enzymes. *Journal of Neurogenetics, 23*(3), 252-261.

Robins, L. N., Babor, T. F., & Cottler, L. B. (1987). *Composite International Diagnostic Interview: Expanded Substance Abuse Module*. St. Louis: Authors.

Robins, L. N., Wing, J., Wittchen, H. U., Helzer, J. E., Babor, T. F., Burke, J., et al. (1988). The Composite International Diagnostic Interview. An epidemiologic Instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry, 45*(12), 1069-1077.

Sareen, J., Bolton, J., Cox, B., & Clara, I. (2006). Use of alcohol and drugs to self-medicate anxiety disorders in a nationally representative sample. *Journal of Nervous and Mental Disease, 194*(11), 818-825.

Schatzkin, A., Abnet, C. C., Cross, A. J., Gunter, M., Pfeiffer, R., Gail, M., et al. (2009).
Mendelian Randomization: How It Can-and Cannot-Help Confirm Causal
Relations between Nutrition and Cancer. *Cancer Prevention Research, 2*(2), 104-
113.

Schumann, G., Coin, L. J., Lourdusamy, A., Charoen, P., Berger, K. H., Stacey, D., et al.
(2011). Genome-wide association and genetic functional studies identify autism
susceptibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption
(vol 108, pg 7119, 2011). *Proceedings of the National Academy of Sciences of the
United States of America, 108*(22), 9316-9316.

Smith, G. D. (2011). Mendelian randomization for strengthening causal inference in
observational studies: Application to gene X environment interactions (vol 5, pg
527, 2010). *Perspectives on Psychological Science, 6*(3), 314-314.

Smith, G. D., & Ebrahim, S. (2003). 'Mendelian randomization': can genetic
epidemiology contribute to understanding environmental determinants of disease?
International Journal of Epidemiology, 32(1), 1-22.

Smith, G. D., & Ebrahim, S. (2005). What can mendelian randomisation tell us about
modifiable behavioural and environmental exposures? *British Medical Journal,
330*, 1076-1079.

Speliotes, E. K., Willer, C. J., Berndt, S. I., Monda, K. L., Thorleifsson, G., Jackson, A.
U., et al. (2010). Association analyses of 249,796 individuals reveal 18 new loci
associated with body mass index. *Nature Genetics, 42*(11), 937-953.

Steinberg, L., Cauffman, E., Woolard, J., Graham, S., & Banich, M. (2009). Are

    Adolescents Less Mature Than Adults? Minors' Access to Abortion, the Juvenile

    Death Penalty, and the Alleged APA "Flip-Flop". *American Psychologist, 64*(7),

    583-594.

Sullivan, P. F., Cichon, S., Craddock, N., Daly, M., Faraone, S. V., Gejman, P. V., et al.

    (2009). A framework for interpreting genome-wide association studies of

    psychiatric disorders The Psychiatric GWAS Consortium Steering Committee.

    Molecular Psychiatry, 14(1), 10-17.

Tanner, J. M., Whitehouse, R. H., & Takaishi, M. (1966). Standards from Birth to

    Maturity for Height, Weight, Height Velocity, and Weight Velocity: British

    Children, 1965. Part II. *Archives of Disease in Childhood, 41*, 613-635.

Tarter, R. E., Vanyukov, M., Kirisci, L., Reynolds, M., & Clark, D. B. (2006). Predictors

    of marijuana use in adolescents before and after licit drug use: Examination of the

    gateway hypothesis. *American Journal of Psychiatry, 163*(12), 2134-2140.

Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M.,

    Koseki, M., et al. (2010). Biological, clinical and population relevance of 95 loci

    for blood lipids. *Nature, 466*(7307), 707-713.

Thomas, D. C. (2010a). Gene-environment-wide association studies: emerging

    approaches. *Nature Reviews Genetics, 11*(4), 259-272.

Thomas, D. C. (2010b). Methods for Investigating Gene-Environment Interactions in

    Candidate Pathway and Genome-Wide Association Studies. *Annual Review of

    Public Health, 31*, 21-36.

Thorgeirsson, T. E., Gudbjartsson, D. F., Surakka, I., Vink, J. M., Amin, N., Geller, F., et al. (2010). Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. *Nature Genetics, 42*(5), 448-135.

Timpson, N. J., Lawlor, D. A., Harbord, R. M., Gaunt, T. R., Day, I. N. M., Palmer, L. J., et al. (2005). C-reactive protein and its role in metabolic syndrome: mendelian randomisation study. *Lancet, 366*(9501), 1954-1959.

Vanyukov, M. M., Tarter, R. E., Kirisci, L., Kirillova, G. P., Maher, B. S., & Clark, D. B. (2003). Liability to substance use disorders: 1. Common mechanisms and manifestations. *Neuroscience and Biobehavioral Reviews, 27*(6), 507-515.

Vink, J. M., Smit, A. B., de Geus, E. J. C., Sullivan, P., Willemsen, G., Hottenga, J. J., et al. (2009). Genome-wide Association Study of Smoking Initiation and Current Smoking. *American Journal of Human Genetics, 84*(3), 367-379.

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *American Journal of Human Genetics, 90*(1), 7-24.

Visscher, P. M., Yang, J. A., & Goddard, M. E. (2010). A Commentary on 'Common SNPs Explain a Large Proportion of the Heritability for Human Height' by Yang et al. (2010). *Twin Research and Human Genetics, 13*(6), 517-524.

Voight, B. F., Scott, L. J., Steinthorsdottir, V., Morris, A. P., Dina, C., Welch, R. P., et al. (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nature Genetics, 42*(7), 579-155.

Vrieze, S.I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). *Psychological Methods, 17,* 228-243.

Vrieze, S. I., Hicks, B. M., McGue, M., & Iacono, W. G. (in press). Genetic Influence on the Co-Occurrence of Alcohol, Marijuana, and Nicotine Dependence Symptoms Declines from Age 14 to 29. *American Journal of Psychiatry.*

Vrieze, S. I., Iacono, W. G., & McGue, M. (2012). Confluence of genes, environment, development, and behavior in a post-GWAS world. *Submitted for publication.*

Vrieze, S. I., McGue, M., Miller, M. B., Legrand, L. N., Schork, N. J., & Iacono, W. G. (2011). An assessment of the individual and collective effects of variants on height using twins and a developmentally informative study design. *PLoS Genetics, 7*(12), e1002413.

Welner, Z., Reich, W., Herjanic, B., Jung, K. G., & Amado, H. (1987). Reliability, validity, and parent-child agreement studies of the Diagnostic Interview for Children and Adolescents (DICA). *Journal of the American Academy of Child and Adolescent Psychiatry, 26*(5), 649-653.

Willer, C. J., Sanna, S., Jackson, A. U., Scuteri, A., Bonnycastle, L. L., Clarke, R., et al. (2008). Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nature Genetics, 40*(2), 161-169.

Yang, J. A., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics, 42*(7), 565-131.

Yang, J. A., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *American Journal of Human Genetics, 88*(1), 76-82.

Yang, J. A., Manolio, T. A., Pasquale, L. R., Boerwinkle, E., Caporaso, N., Cunningham, J. M., et al. (2011). Genome partitioning of genetic variation for complex traits using common SNPs. *Nature Genetics, 43*(6), 519-544.

Young, S. E., Corley, R. P., Stallings, M. C., Rhee, S. H., Crowley, T. J., & Hewitt, J. K. (2002). Substance use, abuse and dependence in adolescence: prevalence, symptom profiles and correlates. *Drug and Alcohol Dependence, 68*(3), 309-322.

Zucker, R. A., Heitzeg, M. M., & Nigg, J. T. (2011). Parsing the Undercontrol/Disinhibition Pathway to Substance Use Disorders: A Multilevel Developmental Problem. *Child Development Perspectives, 5*(4), 248-255.

Zuk, O., Hechter, E., Sunyaev, S. R., & Lander, E. S. (2012). The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of the United States of America*.