UvA-DARE (Digital Academic Repository)

Role of overlapping genetic and environmental factors in the relationship between early adolescent conduct problems and substance use in young adulthood

Verweij, K.J.H.; Creemers, H.E.; Korhonen, T.; Latvala, A.; Dick, D.M.; Rose, R.J.; Huizink, A.C.; Kaprio, J.

Role of overlapping genetic and environmental factors in the relationship between early adolescent conduct problems and substance use in young adulthood

Karin J. H. Verweij[1,2]*, Hanneke E. Creemers[3,4]*, Tellervo Korhonen[4,5,6], Antti Latvala[4,5], Danielle M. Dick[7], Richard J. Rose[4,8], Anja C. Huizink[1] & Jaakko Kaprio[4,5,9]

Department of Clinical Developmental Psychology and EMGO Institute for Health and Care Research, VU University, the Netherlands,[1] Department of Biological Psychology/Netherlands Twin Register, VU University, Amsterdam, the Netherlands,[2] Forensic Child and Youth Care Sciences, University of Amsterdam, Amsterdam, the Netherlands,[3] Department of Public Health, University of Helsinki, Helsinki, Finland,[4] National Institute for Health and Welfare, Helsinki, Finland,[5] Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Joensuu, Finland,[6] Departments of Psychiatry, Psychology and Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA, USA,[7] Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA[8] and Finnish Institute of Molecular Medicine, University of Helsinki, Helsinki, Finland[9]

*Joint first authors

ABSTRACT

Aims To determine (1) the prospective associations of conduct problems during early adolescence with tobacco, alcohol and cannabis use in young adulthood and (2) to what extent these associations are due to overlapping genetic versus environmental influences. **Design** A prospective twin study using biometric twin modelling. **Setting** Finland. **Participants** A total of 1847 Finnish twins (943 males and 904 females) were interviewed in early adolescence, 73% of whom (*n* = 1353, 640 males and 713 females) were retained in young adulthood. **Measurements** Symptom counts of conduct disorder (CD) criteria were obtained from a semi-structured clinical interview in early adolescence [age 14–15 years, mean = 14.2, standard deviation (SD) = 0.15]. Frequency of alcohol, tobacco and cannabis use was obtained from a semi-structured clinical interview in young adulthood (age 19.9–26.6 years, mean = 22.4, SD = 0.7). **Findings** We found modest to moderate phenotypical correlations (*r* = 0.16–0.35) between early adolescent CD symptoms and substance use in young adulthood. In males, the phenotypical correlations of CD symptoms with all three substance use variables are explained largely by overlapping genetic influences. In females, overlapping shared environmental influences predominantly explain the phenotypical correlation between CD symptoms and tobacco and cannabis use. **Conclusions** Conduct disorder symptoms in early adolescence appear to moderately predict substance use in early adulthood. In males, genetic influences seem to be most important in explaining the relationship between conduct disorder symptoms and substance use whereas in females, shared environmental influences seem to be most important.

Keywords Adolescent twins, common environments, conduct problems, genetics, substance use, young adult twins.

INTRODUCTION

Adolescence is a period in which many youth initiate the use of psychoactive substances. To illustrate, between ages 15 and 21 years a sharp increase is seen in the incidence of alcohol, tobacco and cannabis use, with estimates of, respectively, 90, 70 and 25–50% in American and European 21-year-olds [1]. Increased levels of substance use are associated with problems later in life, including poor educational performance, health problems and substance use disorders [2–5].

Early conduct problems, including aggression towards people or animals, destruction of property, deceitfulness or theft and serious violations of rules [6], have been associated with various measures of substance use later in life. For instance, White *et al.* [7] found that higher levels of conduct problems among boys in early adolescence predict higher levels of alcohol and marijuana use from ages 13 to

18. In addition, Elkins *et al.* [8] reported that endorsing conduct symptoms by age 11 predicts initiation of tobacco, alcohol and illicit drug use, and that a diagnosis of conduct disorder (CD) between ages 11 and 14 predicts all three types of substance use disorders by age 18. The established association between conduct problems and substance use may be explained by a common aetiology.

Twin studies represent an adequate method to test for common aetiology by determining the extent to which genetic and environmental factors are shared by two or more traits. For both conduct problems and substance use, the contribution of genetic and environmental factors has been demonstrated. With regard to conduct problems, moderate heritability has been shown, with reported estimates ranging from 0.31 to 0.61 [9–11]. Heritability estimates of substance use have been found to vary across stages of substance use involvement, with the strongest genetic influences for heavier stages of substance use [12,13]. For persistent smoking and life-time cannabis use, meta-analyses point towards moderate heritability, with estimates of 0.46 in females and 0.59 in males for persistent smoking [14] and 0.40 in females and 0.48 in males for cannabis use [13]. For level of alcohol consumption, genetic factors also appear to have a moderate influence, with estimates of approximately 0.50 [15,16].

Previous twin research to understand the phenotypical association between CD and substance use involvement has focused mainly on substance use disorders (e.g. [17–19]). However, studies testing the extent to which the association between conduct problems and level of substance use is due to overlapping genetic or environmental influences are relatively scarce. In a cross-sectional study, Miles *et al.* found that the moderate phenotypical correlation ($r = 0.38$ in males and 0.31 in females) between CD and marijuana use in adolescents was due to overlapping genetic and unique environmental influences on both phenotypes, while shared environmental influences explained only marijuana use [10]. The use of longitudinal data provides the opportunity to also study early manifestations of genetic and environmental influences potentially underlying future substance use. In this way, genetically informative designs can improve our understanding of how the relationship between conduct problems and substance use unfolds over time, elucidating further the mechanisms underlying the development of substance use behaviour. For instance, Korhonen *et al.* [20] found that part of the additive genetic and shared environmental factors underlying externalizing behaviour in early adolescence is also involved in smoking behaviour at age 14 and illicit drug use by age 17. Similarly, Silberg *et al.* [21] demonstrated that early conduct disturbance is related to mid-adolescent substance use through common genetic and shared environmental risk factors.

In this study, we used data on conduct problems and substance use from a large sample of Finnish twins, collected with structured interviews at two time-points covering the crucial period from early adolescence (age 14) to young adulthood (ages 19–26). Detailed information was available on CD symptoms and levels of use of the three most commonly used psychoactive substances: tobacco, alcohol and cannabis. The aims of this study were to determine (1) the prospective associations of conduct problems during early adolescence with tobacco, alcohol and cannabis use in young adulthood and (2) the extent to which these associations were due to overlapping genetic and environmental influences. This information guides further research into the specific genetic and environmental origins of the shared liability for conduct problems and substance use, which may contribute to the development of effective prevention or intervention programmes targeting substance use in adolescence and young adulthood.

METHODS

Participants

The study sample consisted of identical (monozygotic; MZ) and non-identical (dizygotic; DZ) Finnish twins, participating in the FinnTwin12 (FT12) study. FT12 is a population-based twin study that consists of five consecutive birth cohorts (1983–7) [22]. The Central Population Registry was used to identify families with twins. A total of 2724 families (87% of all identified eligible families) returned the family questionnaire, and from these families, 2567 twin pairs completed the baseline questionnaires [23].

For the current study, data were used from an intensive subsample of FT12, consisting of (1) a sample drawn randomly from the 1983 cohort, with the exclusion of those living in the most remote rural areas of Finland (13%), (2) a random sample from the 1984–7 cohorts (59%) and (3) all remaining twins from the 1984–7 cohorts at elevated familial risk for alcohol problems based on parental reports of alcohol use (28%) [17]. Previously it was shown that the modest sample enrichment for familial alcohol use did not affect systematically the genetic and environmental variance components for various behavioural variables, including measures of drinking frequency, smoking initiation and behaviour problems [17]. The intensive subsample included 1035 twin pairs who were invited to complete a full psychiatric face-to-face interview in early adolescence (age 14) and young adulthood (ages 19–26). At age 14, 90% of the twins completed the assessment [17,22] and at age 19–26, 73% of the twins studied at age 14 were retained. Attrition was not selective on early adolescent conduct problems or substance use in early or mid-adolescence (all *P*-values > 0.05).

Measures

Conduct problems

At age 14, face-to-face interviews were conducted using a translated version of the Child Semi-Structured Assessment for the Genetics of Alcoholism, Adolescent version (C-SSAGA-A). This polydiagnostic instrument marks life-time diagnoses on a variety of psychiatric disorders, including CD, based on the criteria of the Diagnostic and Statistical Manual of Mental Disorders (DSM)-IIIR [24]. For this study, symptom counts of 13 CD criteria were categorized into no conduct problems, some conduct problems (endorsement of one or two symptoms) and three or more conduct symptoms.

Tobacco, alcohol and cannabis use

In young adulthood, the translated adult version of the SSAGA, a comprehensive, semi-structured interview to assess substance dependence and related psychiatric disorders, was administered [25]. The first questions of the substance use sections were developed to assess life-time use and frequency of use of tobacco, alcohol and cannabis. For this study, frequency of tobacco use was defined as the number of events per day, including cigarettes, cigars, pipes and snus (a Swedish type of smokeless tobacco). Four categories of daily tobacco use were distinguished, ranging from no daily tobacco use to 11 or more cigarettes, cigars, pipes or snus portions a day. Alcohol use was defined as the typical number of drinks a week, including beer, wine, spirits and other alcoholic drinks. Five categories of alcohol use were distinguished, ranging from no/low levels of alcohol use (fewer than three alcoholic drinks a week) to high levels of use (15 or more drinks a week). Cannabis use

was categorized as no life-time use of cannabis, one to two times used during life-time and life-time cannabis use of three times or more. For each substance, the number of categories was based on the distribution of the frequency of use. Table 1 presents the categories for each substance.

Statistics

Genetic model fitting

We used the classical twin design to separate the variance in conduct symptoms during early adolescence and use of tobacco, alcohol and cannabis in young adulthood into that due to additive genetic (A), shared (family) environmental (C) and residual (E) influences (see [26,27]). A represents the variance resulting from the sum of allelic effects across all segregating genes. C refers to environmental influences that have the effect of making twins more similar to one another, such as shared home environment, parenting style and residential area. The E term includes environmental influences that have the effect of making twins different from one another and stochastic biological effects, as well as measurement error. These variance components can be estimated because MZ twins share all their genetic variation, while DZ twins share, on average, 50% of their segregating genetic variation. If the variances in a phenotype were due completely to genetic variance, we would expect a twin pair correlation of 1.0 for MZ twins and 0.5 for DZ twins. Both MZ and DZ twin pairs share the C influences, so if C were the only source of variance we would expect a twin pair correlation of 1.0 for both MZ and DZ pairs. E factors are unique for each individual, so if E were the only source of variance, we would expect a twin-pair correlation of zero for both MZ and DZ pairs. In reality,

Table 1 Descriptive statistics on the levels of conduct problems in early adolescence and substance use in young adulthood.

Phenotype		Response categories	Prevalence	
			Males	*Females*
Conduct	0	No symptoms	451 (48.2%)	574 (63.7%)
	1	1–2 symptoms	346 (37.0%)	251 (27.9%)
	2	3 or more symptoms	138 (14.8%)	76 (8.4%)
Tobacco	0	No tobacco use	256 (40.6%)	380 (53.8%)
	1	1–5 cigarettes per day	62 (9.8%)	116 (16.4%)
	2	6–10 cigarettes per day	155 (24.6%)	135 (19.1%)
	3	11 or more cigarettes per day	158 (25.0%)	75 (10.6%)
Alcohol	0	0–2.5 drinks per week	89 (14.0%)	188 (26.6%)
	1	3–4.5 drinks per week	72 (11.4%)	131 (18.5%)
	2	5–8.5 drinks per week	93 (14.7%)	232 (32.8%)
	3	9–14.5 drinks per week	189 (29.8%)	123 (17.4%)
	4	15 or more drinks per week	191 (30.1%)	33 (4.7%)
Cannabis	0	Never used	435 (69.0%)	535 (76.1%)
	1	Used 1–2 times in lifetime	87 (13.8%)	79 (11.2%)
	2	Used 3 times or more in life-time	108 (17.1%)	89 (12.7%)

phenotypical variance is due generally to a combination of A, C and E influences and structural equation modelling is used to determine the combination of influences that matches the observed data most accurately.

In the same way, the cross-twin cross-trait correlations (CTCT; e.g. the correlation between conduct problems of twin 1 and cannabis use of twin 2) are used to partition the covariance between conduct symptoms and use of tobacco, alcohol and cannabis into A, C and E influences. When the CTCT correlation is higher for MZ twins than

for DZ twins, this indicates that part of the phenotypical correlation between the two traits is due to overlapping genetic influences. A parallel logic follows for the interpretation of C and E paths that contribute to the covariance between the traits. To model this, we fitted a multivariate Cholesky decomposition with the four phenotypes, in which we estimated A, C and E contributions to each of the variables as well as the A, C and E contributions to the phenotypical covariation between the phenotypes. In this way we obtained a measure of overlap in the genetic
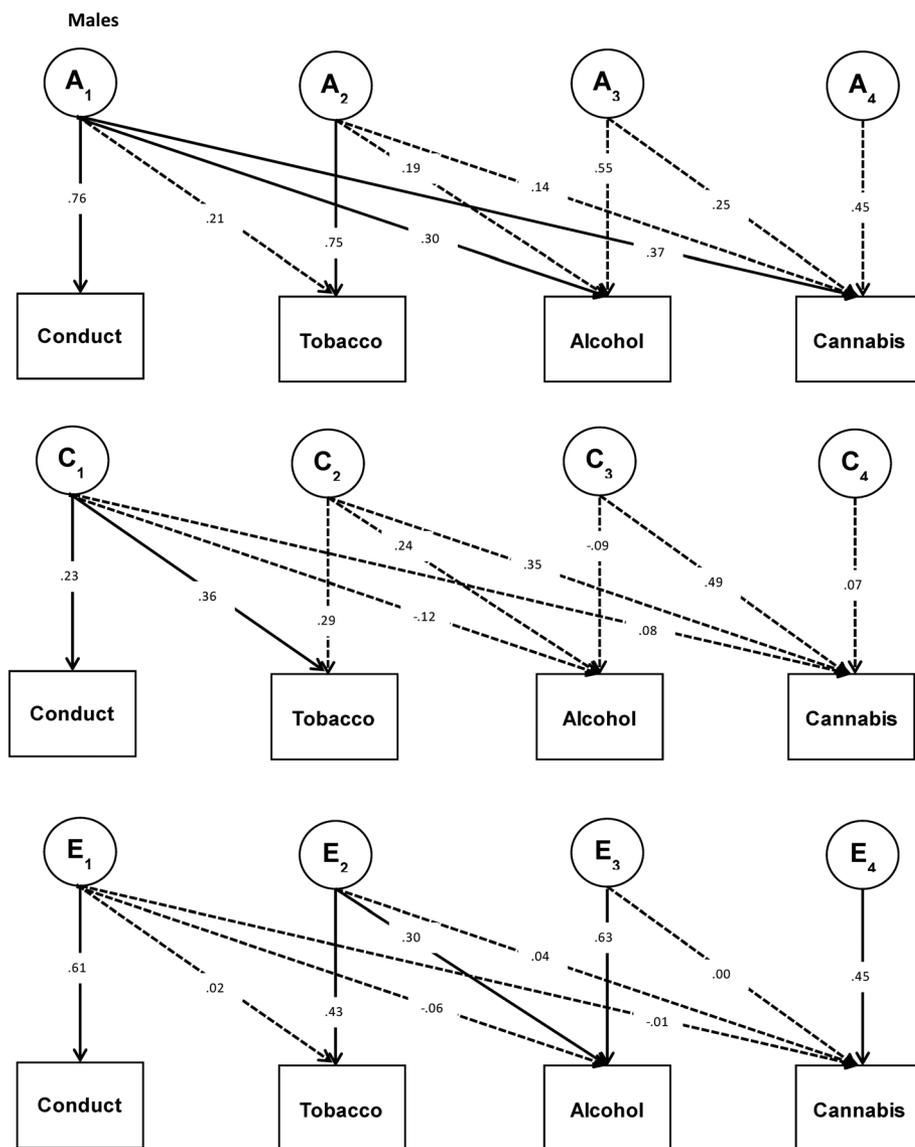

Figure 1 Multivariate Cholesky decomposition. The boxes represent the observed variables and the circles represent the latent variables that influence the observed variables. A, C and E denote latent genetic, shared environmental and residual influences, respectively. A1 represents the genetic influences on CD symptoms, and the cross-paths from A1 to tobacco, alcohol and cannabis use indicate the extent to which these genetic influences are shared with each substance use trait. A2, A3 and A4 represent genetic influences on tobacco, alcohol and cannabis use that are not shared with CD symptoms. Cross-paths from A2 to alcohol and cannabis use and from A3 to cannabis use capture the covariation between these substance use traits. The same structure applies for the C and E factors. Dashed lines indicate non-significant pathways (*P* > 0.05). Note that the pathways from E1 to CD symptoms, E2 to tobacco use, E3 to alcohol use and E4 to cannabis use were not tested for significance; these pathways cannot be dropped from the model, as measurement error is always present
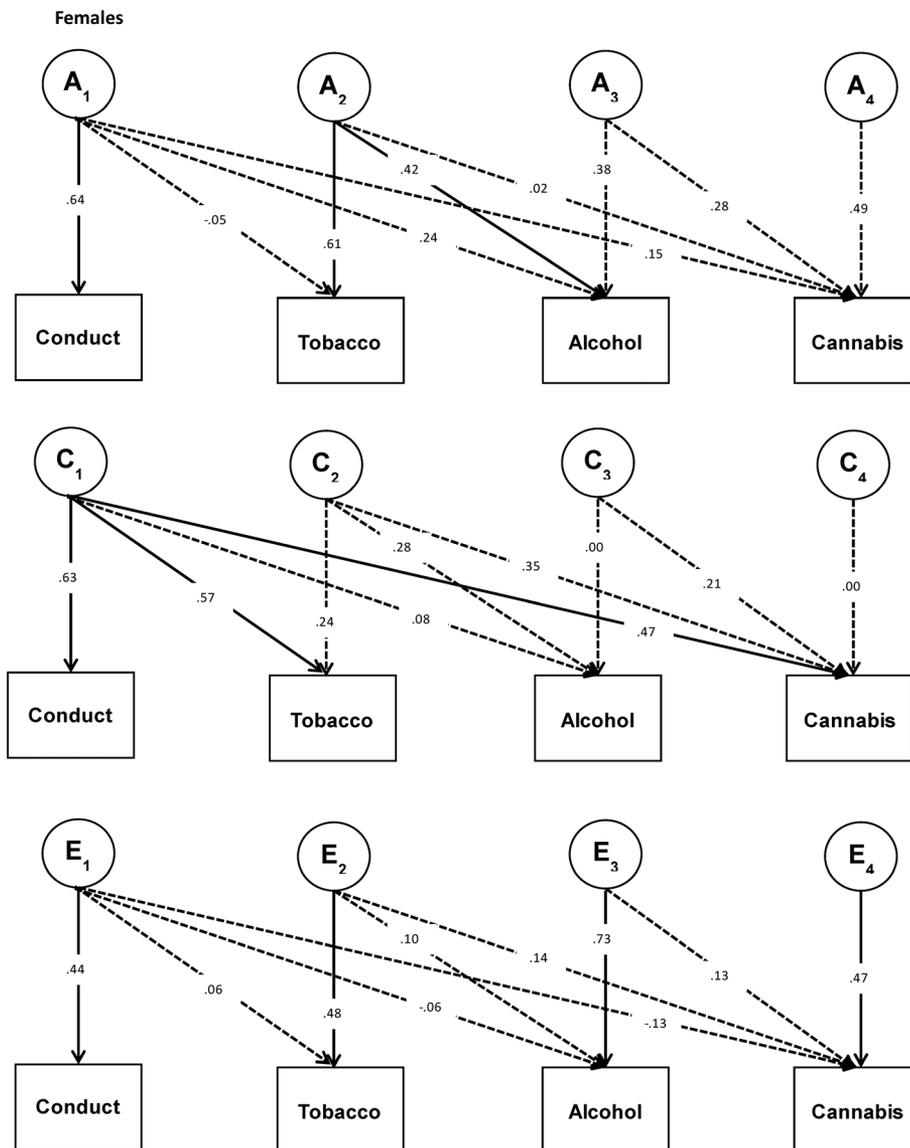
Figure 1 (Continued)

and environmental variation underlying conduct symptoms during adolescence and the use of substances in young adulthood.

Structural equation modelling of twin data is performed in the flexible matrix algebra program Mx [28], which employs maximum likelihood modelling procedures. We fitted models to the raw ordinal data, where it is assumed that a normally distributed continuum of liability underlies the observed categories of the variables. Age and sex effects on the phenotypes were accounted for. The goodness-of-fit of a model is summarized by a statistic distributed as χ^2. By testing the change in model fit ($\Delta\chi^2$) against the change in degrees of freedom (Δdf), we tested whether constraining parameters to zero or constraining them to be equal across groups worsened the model fit significantly. In this way we tested the significance of each pathway in the model.

RESULTS

Descriptive statistics

The final study sample included 1847 twin individuals (943 males and 904 females) with data on at least one variable. Participants were aged between 14.0 and 15.0 years at the time of the first survey [mean = 14.2, standard deviation (SD = 0.15) and between 19.9 and 26.6 at the second survey (mean = 22.4, SD = 0.7)]. At the second survey, 73% of the 1847 twins were retained (N=1353, 640 males and 713 females). The overall sample included 910 complete twin pairs (181 MZ males, 184 MZ females, 152 DZ males, 128 DZ females and 265 DZ opposite-sex pairs) and 27 single twins whose co-twin did not participate. Single twins were included because they contribute to the estimation of the thresholds. Note that sample sizes differ per variable. Table 1 shows the descriptive statistics

for conduct symptoms and each of the substance use variables.

Preliminary analyses

Before estimating the variance components, the effects of sex, age and zygosity on the thresholds of each variable were tested using an α-level of 0.01. Levels of conduct symptoms and substance use did not differ significantly between MZ and DZ twins in either sex. We found a significant sex effect on the thresholds for conduct problems, alcohol and tobacco use (all $P < 0.001$, respectively), indicating that males report more conduct problems and use more alcohol and tobacco than females. No significant sex effects were found for cannabis use, suggesting no differences in levels of cannabis use between sexes. No significant age effects were found for any of the variables. In subsequent genetic analyses thresholds were estimated separately for males and females, and age was included as a covariate.

Table 2 shows the polychoric twin pair correlations for each zygosity group. For each variable, the MZ twin pair correlations were higher than the DZ twin pair correlations, suggesting the influence of genetic factors.

Genetic modelling

We first fitted general sex-limitation models for each of the variables, allowing for qualitative and quantitative differences in the sources of variation between the sexes. Subse-

quently, we fitted common effects sex-limitation models, which only allow for quantitative sex differences in the variance components. Comparing the model fit between these models indicated no significant deterioration in fit for any of the traits (all $P > 0.05$), implying no differences in the source of genetic variation between males and females. We then tested for quantitative sex differences by constraining the variance estimates to be equal across sexes. Male and female path estimates could not be equated for conduct disorder ($\Delta\chi^2_{(2)} = 10.04$, $P = 0.007$). Therefore, in subsequent modelling we estimated the variance components separately for males and females.

Table 3 shows the A, C and E estimates as obtained from the univariate common effects sex-limitation models. Genetic influences on conduct symptoms are 38% for males and 81% for females, whereas genetic influences on the substance use phenotypes range from 31 to 61%. Shared environmental influences play a role in some traits; C explains approximately 20% of the individual difference in conduct symptoms in males, and between 0% (for alcohol use, both sexes) and 47% (for cannabis use in males) in the different substance use variables.

Table 4 shows the phenotypical correlations between conduct symptoms and the three substance use variables as well as the cross-twin cross-trait (CTCT) correlations (e.g. correlation of conduct symptoms of twin 1 with alcohol use of twin 2). We found significant modest to moderate phenotypical correlations between conduct symptoms in early adolescence and all three measures of substance

Table 2 Polychoric twin pair correlations (and 95% confidence intervals) for conduct symptoms and alcohol, cannabis and tobacco use, estimated in Mx (corrected for age and sex effects).

Zygosity	n pairs	Conduct symptoms	Tobacco use	Alcohol use	Cannabis use
MZ M	113–178	0.60 (0.46; 0.71)	0.81 (0.71; 0.88)	0.53 (0.36; 0.66)	0.79 (0.63; 0.89)
DZ M	85–151	0.39 (0.21; 0.55)	0.45 (0.21; 0.64)	0.26 (0.02; 0.46)	0.63 (0.36; 0.80)
MZ F	145–183	0.83 (0.73; 0.89)	0.77 (0.66; 0.84)	0.48 (0.32;0.60)	0.73 (0.52; 0.86)
DZ F	98–128	0.51 (0.30; 0.68)	0.55 (0.33; 0.70)	0.21 (-0.01; 0.41)	0.52 (0.23; 0.73)
DZ OS	168–262	0.33 (0.16; 0.48)	0.55 (0.39; 0.68)	0.29 (0.12; 0.43)	0.47 (0.27; 0.63)

F = female; M = male; MZ = monozygotic; DZ = dizygotic; OS = opposite sex.

Table 3 Estimates of the standardized genetic (A), shared environmental (C) and residual variance (E) components for conduct symptoms and tobacco, alcohol and cannabis use as obtained from univariate twin models (95% confidence intervals between brackets).

Phenotype	Males			Females		
	A	C	E	A	C	E
Conduct symptoms	0.38 (0.00; 0.72)	0.21 (0.00; 0.56)	0.41 (0.28; 0.55)	0.81 (0.25; 0.89)	0.02 (0.00; 0.54)	0.17 (0.11; 0.26)
Tobacco use	0.61 (0.29; 0.83)	0.20 (0.01; 0.49)	0.19 (0.12; 0.29)	0.43 (0.09; 0.76)	0.34 (0.04; 0.64)	0.23 (0.16; 0.33)
Alcohol use	0.51 (0.05; 0.66)	0.02 (0.00; 0.41)	0.47 (0.34; 0.63)	0.47 (0.05; 0.60)	0.01 (0.00; 0.36)	0.52 (0.40; 0.67)
Cannabis use	0.31 (0.00; 0.83)	0.47 (0.00; 0.81)	0.21 (0.11; 0.38)	0.57 (0.00; 0.83)	0.17 (0.00; 0.73)	0.26 (0.14; 0.49)

Table 4 Phenotypical and cross-twin cross-trait (CTCT) correlations of conduct symptoms with tobacco, alcohol and cannabis use (95% confidence intervals between brackets).

	Tobacco use	*Alcohol use*	*Cannabis use*
Phenotypical correlation males	0.25 (0.15; 0.35)	0.18 (0.08; 0.27)	0.29 (0.18; 0.40)
Phenotypical correlation females	0.35 (0.25; 0.45)	0.16 (0.06; 0.25)	0.33 (0.21; 0.44)
CTCT monozygotic males	0.22 (0.10; 0.34)	0.21 (0.08; 0.34)	0.30 (0.16; 0.42)
CTCT dizygotic males	0.22 (0.07; 0.37)	0.09 (−0.07; 0.25)	0.18 (0.00; 0.34)
CTCT monozygotic females	0.32 (0.20; 0.43)	0.16 (0.04; 0.27)	0.37 (0.24; 0.50)
CTCT dizygotic females	0.39 (0.24; 0.52)	0.15 (−0.01; 0.30)	0.39 (0.20; 0.55)
CTCT dizygotic opposite-sex	0.20 (0.07; 0.32)	0.09 (−0.04; 0.21)	0.16 (0.02; 0.30)

use in young adulthood, ranging between 0.16 and 0.35. The CTCT correlations provide a first impression of the extent to which the phenotypical correlations between conduct symptoms and the substance use variables are due to A, C and E (this is modelled formally below). For males, the CTCT correlations appear higher for MZ twins than for DZ twins for conduct-alcohol and conduct-cannabis, suggesting that overlapping genetic influences may explain the phenotypical correlation. The CTCT correlation for conduct-tobacco use for males and all female CTCT correlations appear comparable between MZ and DZ twins, suggesting that the phenotypical correlations between these variables may be more influenced by shared environmental factors.

Path coefficients of the multivariate Cholesky decomposition can be found in Fig. 1 and results are summarized in Table 5. To test the significance of each path in the Cholesky decomposition, genetic and environmental parameters were dropped from the model and model fit was compared with the full model using an α-level of 0.05. Because the parameter estimates here are based on multivariate relationships, the estimates may differ from the estimates obtained in the univariate models and the expectations from the bivariate phenotypical correlations presented in Table 4.

Based on the multivariate Cholesky decomposition, we obtained the contribution of A, C and E influences to the phenotypical correlation between the traits (see Table 5). Following the tracing rules of pathway analysis, the parameters in Fig. 1 can be converted into the A, C and E influences on each trait and the genetic and environmental covariance between traits (also see [29]). For example, the genetic covariance between conduct symptoms and tobacco use can be traced through the path from conduct to A1 times the path from A1 to tobacco (i.e. for males: $0.76 \times 0.21 = 0.16$). The genetic correlation can then also be calculated by dividing the genetic covariance between two traits by the square root of the product of the genetic variances of the two traits.

For males, the phenotypical correlation of conduct symptoms with alcohol and cannabis use can be explained mainly by overlapping A influences, whereas the phenotypical correlation between conduct symptoms and tobacco use is due predominantly to overlapping C influences. For females, the phenotypical correlations between conduct symptoms and tobacco and cannabis use are due predominantly to overlapping C influences. Although not significant, the phenotypical correlation between conduct symptoms and alcohol use in females seems to be explained mainly by A influences.

Table 5 Results from the multivariate Cholesky decomposition of conduct symptoms and tobacco, alcohol and cannabis use. Shown are the standardized genetic (A), shared environmental (C) and residual variance (E) components for each trait as well as the phenotypical correlations (Phen *r*) and the genetic, shared environmental and residual contribution to the phenotypical correlation between the traits.

		Variance components			*Phenotypical correlation with conduct disorder*			
		A	*C*	*E*	*Total Phen r*	*A component of Phen r*	*C component of Phen r*	*E component of Phen r*
Males	Conduct	0.57	0.05	0.38	—	—	—	—
	Tobacco	0.60	0.21	0.19	0.26	0.16	0.08	0.01
	Alcohol	0.43	0.08	0.49	0.17	0.23	−0.03	−0.03
	Cannabis	0.43	0.37	0.20	0.30	0.28	0.02	0.00
Females	Conduct	0.41	0.39	0.19	—	—	—	—
	Tobacco	0.37	0.39	0.24	0.35	−0.03	0.36	0.03
	Alcohol	0.37	0.08	0.54	0.18	0.15	0.05	−0.03
	Cannabis	0.33	0.39	0.27	0.33	0.09	0.30	−0.06

DISCUSSION

Using a prospective, longitudinal twin design, we examined the co-occurrence of early adolescent conduct problems derived from structured face-to-face interviews and frequency of tobacco, alcohol and cannabis use in young adulthood, also obtained from structured interviews. Furthermore, we determined the genetic and environmental influences on the four phenotypes, and assessed the extent to which genetic and environmental factors contribute to the covariance between the phenotypes.

Results of the multivariate analyses showed that, in males and females, both additive genetic effects and shared environmental effects contribute to the variance in adolescent conduct problems and tobacco, alcohol and cannabis use in early adulthood. Because the multivariate models include more information, the estimates obtained from the multivariate analyses should be regarded as superior to the univariate estimates. The heritability estimates for conduct problems based on the multivariate models (57% in males, 41% in females) are roughly similar to previously reported estimates in partly overlapping samples [17,20,30] and in accordance with estimates from previous research [9–11]. Heritability estimates for tobacco (60% in males, 37% in females), alcohol (43% in males, 37% in females) and cannabis use (43% in males, 33% in females) are also broadly comparable to estimates from previous studies [13–16].

Modest to moderate correlations (ranging from 0.16 to 0.35) were found between early conduct problems and tobacco, alcohol and cannabis use in young adulthood. These phenotypical correlations are broadly comparable to earlier established correlations between externalizing/ conduct problems in early adolescence and later tobacco, alcohol and other substance use [20,21]. Moderate correlations between conduct problems and cannabis use have also been reported cross-sectionally [10]. This combination of findings may suggest that the strength of the association between early adolescent behavioural problems and subsequent substance use is not attenuated much over time.

Consistent with Korhonen *et al.* [20], who studied the co-occurrence of early externalizing behaviour based on teacher ratings, and mid-adolescent tobacco and drug use in a partly overlapping sample, we found that common genetic and shared environmental influences explained the co-occurrence between adolescent conduct problems and tobacco, alcohol and cannabis use in young adulthood. More specifically, we found that shared environmental factors contribute to the association between adolescent conduct problems and subsequent tobacco and cannabis use in females, and with tobacco use in males. Silberg *et al.* [21] also indicated the importance of common shared environmental influences for conduct disturbance and substance use in both sexes. They further illustrated that family dysfunction and affiliation with deviant peers are implicated in these shared environmental effects. According to Silberg *et al.*, common genetic influences are also involved in the co-occurrence of conduct problems and substance use in both sexes [21]. This is in contrast to our findings, where only in males can the co-occurrence of conduct symptoms and substance use be explained mainly by overlapping genetic influences. A highly heritable latent phenotype, generally referred to as 'behavioural disinhibition' [31], may be implicated in these shared genetic effects. Based on the findings of the present study, this trait seems of particular importance for understanding conduct problems and substance use in males. Future research is needed to replicate the gender differences in the causes of covariation between early conduct problems and later substance use.

Our findings confirm that the presence of conduct problems in early adolescence is an indicator of liability to substance use later in life. This finding adds to the literature on precursors of substance use phenotypes, including externalizing behaviour [20], early age of alcohol use initiation [32] and early cannabis use [33]. Knowledge on the precursors of substance use can guide the selection of high-risk youths for selective prevention programmes.

Our findings may also inform future genome-wide association (GWA) analyses that will help further elucidation of the complex genetic architecture of conduct behaviour and substance use. The overlapping genetic influences on CD symptoms and substance use (especially in males) indicate that genetic variants identified to play a role in conduct behaviour may also contribute to the genetic architecture of vulnerability to externalizing behaviour in general, and vice versa. Multivariate GWA studies could distinguish between genetic variants that are trait-specific versus those that contribute to externalizing behaviour in general, and this may deepen our understanding of the genetic aetiology of these phenotypes.

Strengths of the present prospective study are the relatively large sample and the use of structured interviews conducted by well-trained staff to assess conduct problems and substance use. A feature that may limit the generalizability of the findings to the general population is the modest oversampling for familial alcoholism risk. However, given that the oversampling had no systematic effect on the estimates of the variance components for several risk-relevant outcomes in early adolescence [17], we do not expect it has affected our results strongly. In addition, although examining levels of use of the various substances as opposed to life-time incidence is more informative in young adulthood, this approach assumes that the same factors influence no versus any use and any versus more frequent use. Multi-stage modelling showed that there is a substantial overlap in genetic and environmental factors influencing earlier and later stages of substance use, but also that there are factors that contribute specifically to later stages of use

[34,35]. Finally, although our data set is relatively large, the power to discriminate between the different variance components was limited due to the use of a threshold model (as necessitated by the highly skewed data) [36].

In conclusion, this study demonstrated that modest to moderate correlations between early adolescent conduct problems and frequency of tobacco, alcohol and cannabis use in early adulthood can be explained by common genetic and shared environmental effects. Our findings point in the direction of sex differences in the causes of covariation between early conduct problems and later substance use frequency. Future research in genetically informative samples is needed to replicate these sex differences. Moreover, as our findings imply that there may be common biological and environmental processes underlying early adolescent conduct problems and later substance use, research directed at the identification of the specific genetic and environmental factors that underlie conduct problems and substance use will improve our understanding of the underlying mechanisms. This may eventually lead to better prevention and treatment interventions.

Declaration of interests

None.

Acknowledgements

References

1. Degenhardt L., Chiu W.-T., Sampson N., Kessler R. C., Anthony J. C., Angermeyer M. *et al.* Toward a global view of alcohol, tobacco, cannabis, and cocaine use: findings from the WHO World Mental Health Surveys. *PLOS Med* 2008; **5**: e141.

2. Brook J. S., Brook D. W., Zhang C., Cohen P. Tobacco Use and Health in Young Adulthood. *J Genet Psychol Res Theory Hum Devel* 2004; **165**: 310–23.

3. King K. M., Meehan B. T., Trim R. S., Chassin L. Marker or mediator? The effects of adolescent substance use on young adult educational attainment. *Addiction* 2006; **101**: 1730–40.

4. McCambridge J., McAlaney J., Rowe R. Adult consequences of late adolescent alcohol consumption: a systematic review of cohort studies. *PLOS Med* 2011; **8**: e1000413. doi: 10.1371/journal.pmed.1000413.

5. Rose RJ, Winter T, Viken RJ, Kaprio J. Adolescent alcohol abuse and adverse adult outcomes: evaluating confounds with drinking-discordant twins. *Alcohol Clin Exp Res* 2014;**38**:2314–21. PMC4146743.

6. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th edn. Arlington, VA: American Psychiatric Publishing; 2013.

7. W. H. R., Zie M., Thompson W., Loeber R., Stouthamer-Loeber M. Psychopathology as a predictor of adolescent drug use trajectories. *Psychol Addict Behav* 2001; **15**: 210–8.

8. Elkins I. J., McGue M., Iacono W. G. Prospective effects of attention-deficit/hyperactivity disorder, conduct disorder, and sex on adolescent substance use and abuse. *Arch Gen Psychiatry* 2007; **64**: 1145–52.

9. Bertoletti E., Michelini G., Moruzzi S., Ferrer G., Ferini-Strambi L., Stazi M. A. *et al.* A general population twin study of conduct problems and the auditory P300 waveform. *J Abnorm Child Psychol* 2014; **42**: 861–9.

10. Miles D. R., van den Bree M. B., Pickens R. W. Sex differences in shared genetic and environmental influences between conduct disorder symptoms and marijuana use in adolescents. *Am J Med Genet* 2002; **114**: 159–68.

11. Rhee S. H., Waldman I. D. Genetic and environmental influences on antisocial behaviour: a meta-analysis of twin and adoption studies. *Psychol Bull* 2002; **128**: 490–529.

12. Rhee S. H., Hewitt J. K., Young S. E., Corley R. P., Crowley T. J., Stallings M. C. Genetic and environmental influences on substance initiation, use, and problem use in adolescents. *Arch Gen Psychiatry* 2003; **60**: 1256–64.

13. V. K. J. H., Zietsch B. P., Lynskey M. T., Medland S. E., Neale M. C., Martin N. G. *et al.* Genetic and environmental influences on cannabis use initiation and problematic use: a meta-analysis of twin studies. *Addiction* 2010; **105**: 417–30.

14. Li M. D., Cheng R., Ma J. Z., Swan G. E. A meta-analysis of estimated genetic and environmental effects on smoking behaviour in male and female adult twins. *Addiction* 2003; **98**: 23–31.

15. Heath A. C., Martin N. G. Genetic influences on alcohol consumption patterns and problem drinking: results from the Australian NH&MRC twin panel follow-up survey. *Ann NY Acad Sci* 1994; **708**: 72–85.

16. Kendler K. S., Schmitt E., Aggen S. H., Prescott C. A. Genetic and environmental influences on alcohol, caffeine, cannabis, and nicotine use from early adolescence to middle adulthood. *Arch Gen Psychiatry* 2008; **65**: 674–82.

17. Rose R. J., Dick D. M., Viken R. J., Pulkkinen L., Nurnberger J. I. Jr., Kaprio J. Genetic and environmental effects on conduct disorder, alcohol dependence symptoms, and their covariation at age 14. *Alcohol Clin Exp Res* 2004; **28**: 1541–8.

18. Slutske W. S., Heath A. C., Dinwiddie S. H., Madden P. A. F., Bucholz K. K., Dunne M. P. *et al.* Common genetic risk factors for conduct disorder and alcohol dependence. *J Abnorm Psychol* 1998; **107**: 363–74.

19. Grant J. D., Lynskey M. T., Madden P. A. F., Nelson E. C., Few L. R., Bucholz K. K. *et al.* The role of conduct disorder in the relationship between alcohol, nicotine and cannabis use disorders. *Psychol Med* 2015; **45**: 3505–15.

20. Korhonen T., Latvala A., Dick D. M., Pulkkinen L., Rose R. J., Kaprio J. *et al.* Genetic and environmental influences underlying externalizing behaviours, cigarette smoking and illicit drug use across adolescence. *Behav Genet* 2012; **42**: 614–25.

21. Silberg J., Rutter M., D'Onofrio B., Eaves L. Genetic and environmental risk factors in adolescent substance use. *J Child Psychol Psychiatry* 2003; **44**: 664–76.

22. Kaprio J., Pulkkinen L., Rose R. J. Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families. *Twin Res Hum Genet* 2002; **5**: 366–71.

23. Dick D. M., Viken R. J., Kaprio J., Pulkkinen L., Rose R. J. Understanding the covariation among childhood externalizing symptoms: genetic and environmental influences on conduct disorder, attention deficit hyperactivity disorder, and oppositional defiant disorder symptoms. *J Abnorm Child Psychol* 2005; **33**: 219–29.

24. Kuperman S., Schlosser S. S., Kramer J. R., Bucholz K. K., Hesselbrock V., Reich T. *et al.* Developmental sequence from disruptive behavior diagnosis to adolescent alcohol dependence. *Am J Psychiatry* 2001; **158**: 2022–6.

25. Bucholz K. K., Cadoret R., Cloninger C. R., Dinwiddie S. H., Hesselbrock V. M., Nurnberger J. I. Jr. *et al.* A new, semistructured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. *J Stud Alcohol* 1994; **55**: 149–58.

26. Neale M. C., Cardon L. R. Methodology for Genetic Studies of Twins and Families. Dordrecht: Kluwer; 1992.

27. Posthuma D., Beem A. L., de Geus E. J. C., van Baal G. C. M., von Hjelmborg J. B., Lachine I. *et al.* Theory and practice in quantitative genetics. *Twin Res* 2003; **6**: 361–76.

28. Neale M. C., Boker S. M., Xie G., Maes H. H. Mx: Statistical Modeling. VCU Box 900126. Richmond, VA: Department of Psychiatry; 2006.

29. Loehlin J. C. The Cholesky approach: a cautionary note. *Behav Genet* 1996; **26**: 65–9.

30. Dibble A. Antisocial Behavior from Adolescence to Early Adulthood: Heritability, Stability, and Correlates using a Longitudinal Twin Sample. VCU Theses and Dissertations; 2013. Available at: http://scholarscompass.vcu.edu/etd/3025 [accessed November 6 2014].

31. Young S. E., Stallings M. C., Corley R. P., Krauter K. S., Hewitt J. K. Genetic and environmental influences on behavioural disinhibition. *Am J Med Genet* 2000; **96**: 684–95.

32. Ystrom E., Kendler K. S., Reichborn-Kjennerud, T. Early age of alcohol initiation is not the cause of alcohol use disorders in adulthood, but is a major indicator of genetic risk. A population-based twin study. *Addiction.* 2014; **109**: 1824–32.

33. Agrawal A., Neale M., Prescott C. A., Kendler K. S. A twin study of early cannabis use and subsequent use and abuse/dependence of other illicit drugs. *Psychol Med* 2004;**34**:1227–37.

34. Agrawal A., Neale M., Jacobson K., Prescott C. A., Kendler K. S. Illicit drug use and abuse/dependence: modeling of two-stage variables using the CCC approach. *Addict Behav* 2005; **30**: 1043–8.

35. Gillespie N. A., Neale M. C., Kendler K. S. Pathways to cannabis abuse: a multi-stage model from cannabis availability, cannabis initiation and progression to abuse. *Addiction* 2009; **104**: 430–8.

36. Neale M. C., Eaves L. J., Kendler K. S. The power of the classical twin study to resolve variation in threshold traits. *Behav Genet* 1994; **24**: 239–58.