Published in final edited form as:

Soc Psychiatry Psychiatr Epidemiol. 2015 August ; 50(8): 1277–1284. doi:10.1007/s00127-015-1030-5.

Environmental clustering of drug abuse in households and communities: multi-level modeling of a national Swedish sample

Kenneth S. Kendler,

Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Box 980126, Richmond, VA 23298-0126, USA. Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA. Department of Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA, USA

Henrik Ohlsson,

Center for Primary Health Care Research, Lund University, Malmö, Sweden

Kristina Sundquist, and

Center for Primary Health Care Research, Lund University, Malmö, Sweden. Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, CA, USA

Jan Sundquist

Center for Primary Health Care Research, Lund University, Malmö, Sweden. Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, CA, USA

Kenneth S. Kendler: kendler@vcu.edu

Abstract

Purpose—Both epidemiological and genetically informative studies indicate that shared environmental influences contribute to resemblance in siblings for drug abuse (DA). To what degree do these influences arise from living in the same household versus residing in the same community?

Methods—We performed a cross-classified multi-level logistic regression on all individuals born in Sweden 1975–1990 ($N = 1558,654$). We assessed the proportion of the total population variation in DA that was due to household versus community effects controlling for genetic resemblance. DA was assessed from medical, criminal and pharmacy records.

Results—Expressed as an intraclass correlation (ICC), the combined household/community effects accounted for ~8 % of the total population variation in DA. The variance attributed to the community was greater than that seen for household (4.5 versus 3.4 %). In males, the variance components were slightly larger and nearly equal at the community (5.3 %) and household level (5.1 %). In females, household effects (4.8 %) were stronger than those arising from the community (3.2 %).

Correspondence to: Kenneth S. Kendler, kendler@vcu.edu.

K. S. Kendler and H. Ohlsson contributed equally to this work.

Conflict of interest None of the authors have conflicts to report.

Conclusion—In the total population and among males, community effects on DA were somewhat more potent than household effects. However, in females, household effects on DA were stronger than community effects. In Sweden, shared environmental effects for DA arise both at the household and at the community level. Community effects on DA are more potent in males than in females.

Keywords

Drug abuse; Environment; Multi-level models; Household; Community; Sweden

Introduction

Drug abuse (DA) is a public health problem of worldwide significance [1]. While prior twin and adoption studies have consistently shown that genetic factors contribute to risk for DA [2–7], three lines of evidence also suggest important environmental influences. First, psychiatric epidemiological studies over the last 75 years have consistently found that rates of DA are strongly related to community-level deprivation, poverty and peer deviance [8–12]. Recent studies have suggested that this association is at least partly causal [13–15]. Second, genetic epidemiological studies including twin [2, 3, 5, 6, 16], adoption [7] and sibling designs [16, 17] have all suggested that “familial” or “shared” environment contributes to DA risk. Third, clustering within neighborhoods and schools has been shown for the use of a range of illicit psychoactive substances [18–20].

The goal of this study is to follow-up on earlier investigations of geographical clustering of psychoactive substance use by studying DA instead of use, comparing environmental effects arising at the level of the household and at the level of the community and, importantly, controlling for the impact of familial/genetic resemblance which has not, to our knowledge, been performed previously in such studies.

Methods

We use linked data from multiple Swedish nationwide registries and healthcare data via the unique individual Swedish ten-digit personal ID number assigned at birth or immigration to all Swedish residents. This ID number was replaced by a serial number to preserve confidentiality.

The following sources were used to create our database: the Total Population Register, containing annual data on family, and geographical status; the Multi-Generation Register, providing information on family relations; the Swedish Hospital Discharge Register, containing all hospitalizations for all Swedish inhabitants from 1964 to 2010; the Swedish Prescribed Drug Register, containing all prescriptions in Sweden picked up by patients from 2005 to 2010; the Outpatient Care Register, containing information from all outpatient clinics from 2001 to 2010; the Primary Health Care Register, containing outpatient primary care data on diagnoses and time for diagnoses 2001–2007 for 1 million patients from Stockholm and middle Sweden; the Swedish Crime Register that included national complete data on all convictions from 1973 to 2011; the Swedish Suspicion Register that included national complete data on all individuals strongly suspected of crime from 1998 to 2011; the

Swedish Mortality Register, containing causes of death; the Population and Housing Censuses that provided information on household, education and geographical status in 1960, 1965, 1970, 1975, 1980, and 1985. A household refers to the person or group of persons who lived in the same dwelling so that we could distinguish individuals living in the same apartment in a multi-apartment dwelling. Each household has its own ID number. Geographical status was defined as Small Areas for Market Statistics (SAMS) that are small geographical units defined by Statistics Sweden, the Swedish government-owned statistics bureau. There are 8766 SAMS throughout Sweden, with a mean (SD) of population size of 938 (1141) in 1990. For the age range we are examining here, the mean (SD) sample size per SAMS in 1990 was 138 (155). For more details on SAMS areas, see http://www.scb.se/Grupp/Produkter_Tjanster/Skraddarsydd/Regionala_produkter/Marknadsprofil/SamsAtlas/index.html#. We secured ethical approval for this study from the Regional Ethical Review Board of Lund University (No. 2008/409).

DA was identified in the Swedish medical registries by ICD codes (ICD8: Drug dependence (304); ICD9: Drug psychoses (292) and Drug dependence (304); ICD10: Mental and behavioral disorders due to psychoactive substance use (F10-F19), except those due to alcohol (F10) or tobacco (F17)); in the Suspicion Register by codes 3070, 5010, 5011, and 5012, that reflect crimes related to DA; and in the Crime Register by references to laws covering narcotics (law 1968:64, paragraph 1, point 6) and drug-related driving offenses (law 1951:649, paragraph 4, subsection 2 and paragraph 4A, subsection 2). DA was identified in individuals (excluding those suffering from cancer) in the Prescribed Drug Register who had retrieved (in average) more than four defined daily doses a day for 12 months from either of hypnotics and sedatives (Anatomical Therapeutic Chemical (ATC) Classification System N05C and N05BA) or Opioids (ATC: N02A).

Study population and outcome variable

Our baseline population includes all individuals born in Sweden between 1975 and 1990, residing in Sweden on 31 December 1990 ($N = 1558,654$). We exclude individuals without mother, father, paternal/maternal grandparents, household or SAMS information included in the registers ($N = 349,216$). The outcome measure that we examine on an individual level is DA that is measured during the period from 1 January 1991 to 31 December 2011 and considered binary.

In the database, we categorize individuals into several different classifications: household, SAMS, ID based on mother and father (defines full-siblings), ID based on mother (defines full and maternal half-siblings), ID based on father (defines full and paternal half-siblings), and IDs based on same maternal grandmother, maternal grandfather, paternal grandmother, and paternal grandfather (all of which define individuals who are at least first cousins).

To study the importance of household and SAMS for understanding individual variation in DA, we used cross-classified multi-level logistic regression. Such a model comprises individuals nested within a cross-classification of two or more different hierarchies. The cross-classified model enables us to take into account influences on the total individual variance coming from different classifications. To simplify the models, we excluded MZ twins, first cousins and paternal half-siblings who shared the same household or the same

SAMS ($N = 119,498$), leaving a sample size of 1089,940. This was necessary to produce a tractable model that included four different classifications for each individual: household, SAMS, ID based on mother, and ID based on mother and father. Because we excluded MZ twins, paternal half-siblings and first cousins, all the genetic effects that could contribute to the variance components for households and SAMS are captured by the two other classifications: ID based on mother (which will identify maternal half-siblings) and ID based on mother and father (which will identify full siblings). We conduct one model for the entire sample ($N = 1089,940$), and one model each for males ($N = 549,915$) and females ($N = 520,216$).

We present the variance components attributed to the different classifications together with a 95 % credible interval (CI). Measures of variance and clustering allow us to identify the environmental level on which contextual influences operate on DA risk. For example, the greater the degree of resemblance for DA status of individuals within the same household (controlling for genetic resemblance), the more likely environmental factors shared by individuals with the same household impact on liability to DA.

We also present the intraclass correlation (ICC) that can be interpreted as the proportion of the total residual variation in DA that is due to differences between classifications. To calculate the ICC, we use the latent variable method [21, 22], which assumes the existence of a latent individual variable that follows a logistic distribution with a variance equal to 3.29 ($\pi^2/3$). The ICC is then only a function of the variance due to the classifications and does not directly depend on the prevalence of DA [23]. The ICC is, therefore, measured on the logistic scale and the cluster and individuals effects are assumed additive. We present three ICCs: ICC_{SAMS} , $ICC_{Household}$ and $ICC_{SAMS + Households}$. The ICC_{SAMS} tells us how much of the individual variation in DA that can be attributed to the SAMS level (when genetic and household effects are taken into account) and it is calculated according to the following formula:

$$\frac{(\text{Variance}_{SAMS})}{(\text{Variance}_{individual}) + (\text{Variance}_{Household}) + (\text{Variance}_{SAMS}) + (\text{Variance}_{MotherID}) + (\text{Variance}_{Mother\&FatherID})}.$$

The $ICC_{Household}$ tells us how much of the individual variation in DA that can be attributed to the household level (when genetic and SAMS effects are taken into account) and it is calculated according to the following formula:

$$\frac{(\text{Variance}_{Household})}{(\text{Variance}_{individual}) + (\text{Variance}_{Household}) + (\text{Variance}_{SAMS}) + (\text{Variance}_{MotherID}) + (\text{Variance}_{Mother\&FatherID})}.$$

The $ICC_{SAMS + Households}$ tells us how much of the individual variation in DA that can be attributed to the household and SAMS level (when genetic effects are taken into account) and it is calculated according to the following formula:

$$\frac{(\text{Variance}_{Household}) + (\text{Variance}_{SAMS})}{(\text{Variance}_{individual}) + (\text{Variance}_{Household}) + (\text{Variance}_{SAMS}) + (\text{Variance}_{MotherID}) + (\text{Variance}_{Mother\&FatherID})}.$$

All analyses are performed using the WinBugs software [24] and the MCMC procedure in MLwiN software [25]. We store the results from each step in the iteration procedure, which gives us a distribution of the parameters of interest, and from this distribution we calculate the median and corresponding 95 % CI.

Results

In the full sample ($N = 1558,654$), 4.0 % of the individuals were registered for DA (5.8 % in males and 2.2 % in females). The rates were identical for the more restricted sample we use here for analysis ($N = 1089,940$). The mean percentages of DA in the 8766 SAMS areas of Sweden and the 636,202 households were 4.0 (SD 4.6) and 4.1 (SD 17.2), respectively.

Table 1 provides the tetrachoric correlations for DA of individuals as a function of their genetic relationship and cohabitation history. Unrelated individuals residing in the same household or the same SAMS but not the same household had very similar and modest correlations for DA ($\sim +0.05$). Among full and maternal half-siblings, resemblance was higher for pairs residing in the same household compared to those living in different SAMS. (The numbers of full and maternal half-siblings living in the same SAMS but different households were too small to produce stable estimates).

Table 1 also presents results separately for male–male and female–female pairs. For male–male pairs, the correlation for non-related individuals in the same SAMS is modestly higher than the correlation for non-related individuals in the same household. In female–female pairs, no difference is seen in correlation between these types of pairs. In male–male full and maternal half-sibling pairs and in female–female full-sibling pairs, resemblance for DA was greater for those living in the same household than in different SAMS. In maternal half-siblings, the reverse pattern is seen but these estimates are known imprecisely.

Table 2 illustrates the results from the cross-classified multi-level logistic model. The variance in DA at the level of SAMS is somewhat greater than the variance at the level of the household. Specifically, of the total variance attributable to SAMS and household, SAMS contributes 57 % and household 43 %. When translating the variances into the ICC, the combined household/SAMS variance accounts for 8 % of the total individual variation in DA. Of note, the genetic part of the model (Variance MotherID and Variance Mother&FatherID) accounts for 4.4 times as much variation than households.

For males, the variance for SAMS is similar to the variance for households. Of the total variation that is attributed to SAMS and household, SAMS contributes 51 % and household 49 %. When translating the variances into the ICC, the combined household/SAMS variance accounts for 10.4 % of the total individual variation in DA. The genetic part of the model accounts for 3.7 times as much variation as households.

For females, the variance for households is higher than the variance for SAMS, even though their 95 % CIs overlapped. Of the total variation that is attributed to SAMS and household, SAMS contributed 40 % and household 60 %. When translating the variances into the ICC, the combined household/SAMS variance accounted for 8 % of the total individual variation

in DA. The genetic part of the model accounts for 4.4 times as much variation than households.

We replicated our results excluding maternal half-siblings instead of paternal half-siblings living in the same household and same SAMS. The results were very similar. In these models, the variance components for SAMS and household are 0.25 and 0.21, respectively (versus 0.26 and 0.20 excluding paternal half-siblings). We also replicated our results in a sample using only individuals born over a 5-year interval (1985–1990) rather than a 15-year interval (1975–1990). Results were again very similar. For example, in these models, the variance components for SAMS and household were 0.25 and 0.20, respectively.

Discussion

The goal of this report was to assess the relative influences on the clustering of cases of DA of living in the same households versus residing in the same community. To address this question, we developed a modified multi-level model that permitted us to examine household and community effects, controlling for the impact of genetic effects on resemblance for DA.

We found in the total Swedish population, born 1975–1990, substantial effects on the clustering of cases of DA, controlling for genetic effects, of both residence in the same household and living in the same SAMS area. In both the total population and among males, community effects on DA were modestly stronger than household effects. However, in females, household effects on DA were stronger than community effects. We would conclude from these findings that in Sweden, shared environmental effects for DA arise both at the household and at the community level. Furthermore, community effects on DA are more potent in males than in females.

These results are consistent with several prior lines of evidence. We are aware of one prior twin study that examined both familial–environmental and school-based effects on substance use in adolescent Finnish twins [26]. Controlling for genetic effects, school-based effects were substantial for both drinking and smoking. For smoking and regular alcohol consumption, family effects were moderately stronger than school effects. However, for drinking with friends without adult supervision, school effects were stronger than familial–environmental effects [26].

Several lines of evidence suggest that compared to females, males are more motivated to use psychoactive substances to conform to subgroup values and are more influenced by peers in their intake of drugs [27, 28]. In a relevant Swedish high-school student survey, Svensson [29] found that males had higher exposure rates to deviant peers than females because parents of girls monitored their behavior and friends more closely than did the parents of boys. He also found that the probability of drug use was more strongly predicted by peer deviance in males than in females [29]. Both more frequent exposure to peer deviance and a greater impact of that exposure on drug use in males versus females would likely translate into stronger community influences on DA risk in young men than in young women.

Our present findings are also congruent with three of our previous efforts to address the question of shared environmental influences on DA in Sweden. First, we found that resemblance for DA in sibling pairs is greater for those closer to one another in age [17]. These findings likely result, at least in part, from greater co-socialization with community peers in closer aged siblings. Second, using standard twin modeling, we found stronger evidence for shared environmental effects of DA in males versus females [16]. Third, we found in male–male full- and half-sibling pairs, resemblance in risk for DA was predicted by years of living together in the same household and to a slightly weaker extent by years of residing in the same community but not the same household. We then examined the same effects in female–female half- and full-sibling pairs and found that they were present but weaker than those observed in the male–male pairs [16].

Limitations

These results should be interpreted in the context of six potentially important methodological limitations. First, we detected subjects with DA from medical, legal and pharmacy records, a method that does not require accurate respondent recall. However, such data likely contain both false negative and false positive diagnoses, the frequency of which we cannot easily estimate. An epidemiological study of DA conducted in neighboring Norway, which has similar rates of drug use and abuse to Sweden [30, 31], found lifetime prevalence rates of DSM-III-R [32] DA similar to those found using our registry based methods [33]. Marked under-ascertainment of DA in our sample is, therefore, unlikely. Furthermore, the validity of our assessment methods for DA is supported by the very high ORs (mean 2.2 [7]) for registration for DA across our different registry sources.

Second, while the household is a naturally defined unit, the geographical boundaries of the SAMS, while generally set to reflect natural community boundaries, were to some extent arbitrary and set for administrative reasons. It is possible that other criteria, such as municipalities, school districts, or geographic units centered on individuals' residences, might have done a superior job of capturing community-level influences. However, despite what might be less than ideal boundaries, the SAMS units in our analyses indeed captured a substantial collective influence on risk for DA. Furthermore, the interpretation of variance is constrained by time and geographical location [34–36], so that the results only pertain to the population in Sweden in 1990.

Third, we cannot rule out the possibility that the household/SAMS effect might vary for different types of individuals [37]. But as the information at the household level was rather sparse, a further stratification beyond gender would likely be uninformative. The comparison between males and females might also have been more appropriate within the same model. However, due to convergence problems, we chose to analyze results for males and females separately.

Fourth, we could not rule out that clustering of DA within SAMS areas might be influenced by extraneous factors such as differences in rates of conviction for drug-related offenses or in diagnostic practices of local physicians. To partly examine this possible bias, we repeated the analyses presented removing cases of DA identified only through the crime register. The

combined ICC for household and SAMS remained high (8 %), suggesting that the clustering we present could not arise due to police practice.

Fifth, might some of the concentration of DA within SAMS areas result from self-selection of high risk families (e.g., positive gene–environmental correlation)? If this was the case, we would expect that those moving into a SAMS should have similar rates of DA to those that already live in the area. That is, DA should not cluster more strongly in families residing for a longer period of time in a SAMS. To evaluate this hypothesis, we estimated a model for the parents of the probands in our sample and investigated if the clustering of DA was higher for those residing at least 5 years in the area compared to those who had resided less than 5 years. The results were opposite to what would be expected if self-selection had a major influence on our results. The ICC for DA in parents for those who had lived 5 years or more was as high as 12 %, while for those who had lived in the area for less than 5 years was only 7 %.

Finally, our analysis did not account for age composition of the households and SAMS areas. As a result, it is possible that residual compositional confounding [38] might explain the variance we detected in our study. It should also be noted that the information on households and SAMS are measured when the individuals are between 0 and 15 years old, while DA is measured when the individual is 15 years or older, suggesting that the follow-up time is between 6 and 21 years. However, using only individuals born over a 5-year interval produced very similar results. It should also be noted that the SAMS effect was attenuated the longer the follow-up; the ICC for SAMS when using DA registration during 1990–2000 as outcome was 11 %, while the ICC for SAMS using DA registration during 2001–2011 as outcome was 7 %. This can be both due to the fact that the prevalence of DA increased over time [39], but also due to the fact that individuals could move during follow-up. Nevertheless, the ICC for SAMS was still strong, suggesting noticeable clustering of DA within SAMS areas.

Conclusion

Using a statistical method that controlled for genetic effects, we showed evidence for appreciable clustering of DA cases in Sweden both within households and within small communities. These results suggest that several distinct environmental mechanisms underlie shared environmental effects on risk for DA that have been previously detected in Sweden using more conventional twin and adoption methods in Sweden [7, 16]. Some of these, such as premature loss of parents, the quality of parent–child relationships, parental monitoring, sibling effects, and religiosity—all of which have been shown to impact on risk for DA [7, 40–43]—likely act largely at the level of the household. Other important influences, including drug availability and peer effects—which also impact substantially on drug use and abuse [44, 45]—are probably operating at the level of the community. As we have argued elsewhere [16], just as molecular genetic methods are now being used to clarify, at a more mechanistic level, the impact of genetic risk factors for psychiatric and drug abuse disorders, social epidemiological analytic tools need to be applied to aggregate environmental effects to clarify at a more specific level their mode of action. Finally, the statistical tools we developed here have wide applicability to other disorders and similar

epidemiological data sets. It would be of particular interest to determine, for a range of psychiatric and substance use disorders, how they differ in the proportion of variance arising from household and community effects.

Acknowledgments

This study was funded by grant RO1 DA030005 from the National Institute of Drug Abuse, the Swedish Research Council 2008-3110, 2008-2638, the Swedish Research Council for Health, Working Life and Welfare (Reg.nr: 2013-1836), the ALF project grant, Lund, Sweden and the Swedish Council for Information on Alcohol and Other Drugs (CAN).

References

1. World Health Organization. Atlas on substance use (2010) Resources for the prevention and treatment of substance use disorders. 2010 [NLM classification: WM 270]. WHO Library Cataloguing-in-Publication Data. WHO Press;
2. Tsuang MT, Lyons MJ, Eisen SA, Goldberg J, True W, Lin N, Meyer JM, Toomey R, Faraone SV, Eaves L. Genetic influences on DSM-III-R drug abuse and dependence: a study of 3,372 twin pairs. *Am J Med Genet.* 1996; 67:473–477. [PubMed: 8886164]
3. van den Bree MB, Johnson EO, Neale MC, Pickens RW. Genetic and environmental influences on drug use and abuse/dependence in male and female twins. *Drug Alcohol Depend.* 1998; 52:231–241. [PubMed: 9839149]
4. Kendler KS, Prescott CA. Cannabis use, abuse, and dependence in a population-based sample of female twins. *AJP.* 1998; 155:1016–1022.
5. Lynskey MT, Heath AC, Nelson EC, Bucholz KK, Madden PA, Slutske WS, Statham DJ, Martin NG. Genetic and environmental contributions to cannabis dependence in a national young adult twin sample. *Psychol Med.* 2002; 32:195–207. [PubMed: 11871371]
6. Kendler KS, Karkowski LM, Neale MC, Prescott CA. Illicit psychoactive substance use, heavy use, abuse, and dependence in a US population-based sample of male twins. *Arch Gen Psychiatry.* 2000; 57:261–269. [PubMed: 10711912]
7. Kendler KS, Sundquist K, Ohlsson H, Palmer K, Maes H, Winkleby MA, Sundquist J. Genetic and familial environmental influences on the risk for drug abuse: a national Swedish adoption study. *Arch Gen Psychiatry.* 2012; 69:690–697. [PubMed: 22393206]
8. Faris, RE.; Dunham, HW. Mental disorders in urban areas: an ecological study of schizophrenia and other psychoses. University of Chicago Press; Chicago: 1939.
9. Compton WM, Thomas YF, Stinson FS, Grant BF. Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: results from the national epidemiologic survey on alcohol and related conditions. *Arch Gen Psychiatry.* 2007; 64:566–576. [PubMed: 17485608]
10. Muntaner C, Eaton WW, Dula C, Kessler RC, Sorlie PD. Social class, assets, organizational control and the prevalence of common groups of psychiatric disorders. *Soc Sci Med.* 1998; 47:2043–2053. [PubMed: 10075245]
11. Warner LA, Kessler RC, Hughes M, Anthony JC, Nelson CB. Prevalence and correlates of drug use and dependence in the United States. Results from the national comorbidity survey. *Arch Gen Psychiatry.* 1995; 52:219–229. [PubMed: 7872850]
12. Kendler KS, Ohlsson H, Sundquist K, Sundquist J. Peer deviance, parental divorce, and genetic risk in the prediction of drug abuse in a nationwide Swedish sample: evidence of environment-environment and gene-environment interaction. *JAMA Psychiatry.* 2014; 71:439–445. [PubMed: 24576925]
13. Dohrenwend BP, Levav I, Shrout PE, Schwartz S, Naveh G, Link BG, Skodol AE, Stueve A. Socioeconomic status and psychiatric disorders: the causation-selection issue. *Science.* 1992; 255:946–952. [PubMed: 1546291]

14. Sareen J, Afifi TO, McMillan KA, Asmundson GJ. Relationship between household income and mental disorders: findings from a population-based longitudinal study. *Arch Gen Psychiatry*. 2011; 68:419–427. [PubMed: 21464366]
15. Kendler KS, Ohlsson H, Sundquist K, Sundquist J. The causal nature of the association between neighborhood deprivation and drug abuse: a prospective national Swedish co-relative control study. *Psychol Med*. 2014; 44(12):2537–2546. [PubMed: 25055172]
16. Kendler KS, Maes HH, Sundquist K, Ohlsson H, Sundquist J. Genetic and family and community environmental effects on drug abuse in adolescence: a Swedish national twin and sibling study. *AJP*. 2013; 171:209–217.
17. Kendler KS, Ohlsson H, Sundquist K, Sundquist J. Within-family environmental transmission of drug abuse: a Swedish national study. *JAMA Psychiatry*. 2013; 70:235–242. [PubMed: 23229904]
18. Bobashev GV, Anthony JC. Clusters of marijuana use in the United States. *Am J Epidemiol*. 1998; 148:1168–1174. [PubMed: 9867262]
19. Petronis KR, Anthony JC. Perceived risk of cocaine use and experience with cocaine: do they cluster within US neighborhoods and cities? *Drug Alcohol Depend*. 2000; 57:183–192. [PubMed: 10661669]
20. Delva J, Bobashev G, Gonzalez G, Cedeno M, Anthony JC. Clusters of drug involvement in Panama: results from Panama's 1996 national youth survey. *Drug Alcohol Depend*. 2000; 60:251–257. [PubMed: 11053759]
21. Goldstein, H. Multilevel statistical models. 3. Hodder Arnold; London: 2003.
22. Goldstein H, Browne W, Rasbash J. Partitioning variation in generalised linear multilevel models. *Underst Stat*. 2002; 1:223–232.
23. Wu S, Crespi CM, Wong WK. Comparison of methods for estimating the intraclass correlation coefficient for binary responses in cancer prevention cluster randomized trials. *Contemp Clin Trials*. 2012; 33:869–880. [PubMed: 22627076]
24. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput*. 2000; 10:325–337.
25. Rasbash, J.; Charlton, C.; Brown, WJ.; Healy, M.; Cameron, B. MCMC estimation in MLwiN [computer program]. Version 2.1 Centre for multilevel modelling. University of Bristol; 2009.
26. Rose RJ, Viken RJ, Dick DM, Bates JE, Pulkkinen L, Kaprio J. It does take a village: nonfamilial environments and children's behavior. *Psychol Sci*. 2003; 14:273–277. [PubMed: 12741753]
27. Graziano F, Bina M, Giannotta F, Ciairano S. Drinking motives and alcoholic beverage preferences among Italian adolescents. *J Adolesc*. 2012; 35:823–831. [PubMed: 22240325]
28. Borsari B, Carey KB. Peer influences on college drinking: a review of the research. *J Subst Abuse*. 2001; 13:391–424. [PubMed: 11775073]
29. Svensson R. Gender differences in adolescent drug use—the impact of parental monitoring and peer deviance. *Youth Soc*. 2003; 34:300–329.
30. Kraus L, Augustin R, Frischer M, Kummeler P, Uhl A, Wiessing L. Estimating prevalence of problem drug use at national level in countries of the European Union and Norway. *Addiction*. 2003; 98:471–485. [PubMed: 12653817]
31. Hibell, B.; Guttormsson, U.; Ahlstrom, S.; Balakireva, O.; Bjarnason, T.; Kokkevi, A.; Kraus, L. The 2007 ESPAD report: substance use among students in 35 European countries. The Swedish Council for Information on Alcohol and Other Drugs (CAN); Stockholm, Sweden: 2007.
32. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, revised. 3. American Psychiatric Association; Washington: 1987.
33. Kringlen E, Torgersen S, Cramer V. A Norwegian psychiatric epidemiological study. *AJP*. 2001; 158:1091–1098. [PubMed: 11238057]
34. Braumoeller B. Explaining variance: or, stuck in a moment we can't get out of. *Polit Anal*. 2006; 14:268–290.
35. Downs G, Rocke D. Interpreting heteroscedasticity. *Am J Polit Sci*. 1979; 23:816–828.
36. Lewontin RC. The analysis of variance and the analysis of causes. 1974. *Int J Epidemiol*. 2006; 35:520–525. [PubMed: 16645033]

37. Chaix B, Rosvall M, Merlo J. Neighborhood socioeconomic deprivation and residential instability: effects on incidence of ischemic heart disease and survival after myocardial infarction. *Epidemiology*. 2007; 18:104–111. [PubMed: 17130687]
38. Oakes JM. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Soc Sci Med*. 2004; 58:1929–1952. [PubMed: 15020009]
39. Giordano GN, Ohlsson H, Kendler KS, Winkleby MA, Sundquist K, Sundquist J. Age, period and cohort trends in drug abuse hospitalizations within the total Swedish population (1975–2010). *Drug Alcohol Depend*. 2013; 134:355–361. [PubMed: 24300899]
40. Kendler KS, Myers J, Prescott CA. Parenting and adult mood, anxiety and substance use disorders in female twins: an epidemiological, multi-informant, retrospective study. *Psychol Med*. 2000; 30(2):281–294. [PubMed: 10824649]
41. Kendler KS, Neale MC, Prescott CA, Kessler RC, Heath AC, Corey LA, Eaves LJ. Childhood parental loss and alcoholism in women: a causal analysis using a twin-family design. *Psychol Med*. 1996; 26:79–95. [PubMed: 8643766]
42. Chilcoat HD, Dishion TJ, Anthony JC. Parent monitoring and the incidence of drug sampling in urban elementary school children. *Am J Epidemiol*. 1995; 141:25–31. [PubMed: 7801962]
43. Brook JS, Whiteman M, Gordon AS, Brook DW. The role of older brothers in younger brothers' drug use viewed in the context of parent and peer influences. *JGP*. 1990; 151:59–75.
44. Gatti U, Tremblay RE, Vitaro F, McDuff P. Youth gangs, delinquency and drug use: a test of the selection, facilitation, and enhancement hypotheses. *J Child Psychol Psychiatr*. 2005; 46:1178–1190. [PubMed: 16238665]
45. van den Bree MB, Pickworth WB. Risk factors predicting changes in marijuana involvement in teenagers. *Arch Gen Psychiatry*. 2005; 62:311–319. [PubMed: 15753244]

Table 1
Tetrachoric correlations among all possible pairs and same-sex pairs in the Swedish population born 1975–1990

	Male-male		Female-female	
	Correlation (SE)	N	Correlation (SE)	N
All				
Unrelated individuals				
SHH	0.044 (0.02)	15,808	0.035 (0.04)	4278
SSAMS	0.052 (0.00)	6.53×10^8	0.073 (0.00)	1.73×10^8
Full-siblings				
SHH	0.38 (0.00)	1198,300	0.47 (0.01)	319,924
DSAMS	0.34 (0.03)	10,884	0.41 (0.04)	2936
Maternal half-siblings				
SHH	0.25 (0.01)	95,462	0.33 (0.01)	24,014
DSAMS	0.24 (0.02)	14,079	0.23 (0.03)	4276

SHH lived in the same household, SSAMS lived in the same SAMS, DSAMS lived in the different SAMS

Table 2

Results from the cross-classified logistic regression model with 95 % credible intervals

	All	Males	Females
Variance SAMS	0.26 (0.24; 0.28)	0.35 (0.32; 0.38)	0.19 (0.16; 0.22)
Variance Household	0.20 (0.17; 0.22)	0.33 (0.22; 0.43)	0.28 (0.19; 0.41)
Variance MotherID	1.32 (1.16; 1.44)	1.77 (1.33; 1.97)	1.18 (0.85; 1.60)
Variance Mother&FatherID	0.68 (0.57; 0.87)	0.75 (0.66; 1.44)	0.87 (0.55; 1.19)
Intraclass correlation ^a			
ICC Household	3.4 % (3.0; 3.9)	5.1 % (3.3; 6.6)	4.8 % (3.3; 6.8)
ICC SAMS	4.5 % (4.2; 4.8)	5.3 % (4.9; 5.7)	3.2 % (2.7; 3.7)
ICC SAMS + Household	8.0 % (7.4; 8.5)	10.4 % (8.6; 11.9)	8.0 % (6.3; 10.1)

^aThe intraclass correlation (ICC) is a function of the variance components in the analysis as shown by the formulas above. In order to calculate the 95 % credible intervals we stored the results from each step in the iteration procedure (150,000 iterations after a burn-in phase of 1000 iterations), which gave us a distribution of the parameters of interest, and from this distribution, we calculated the ICC at each step of the iteration procedure and we present the median value and corresponding 95 % CI