2017-03-02

# Cognitive ability and risk for substance misuse in men : genetic and environmental correlations in a longitudinal nation-wide family study

Latvala, Antti; Kuja-Halkola, Ralf; D'Onofrio, Brian M; Larsson, Henrik; Lichtenstein, Paul

Addiction. 2016 Oct;111(10):1814-22.
http://doi.org/10.1111/add.13440
http://hdl.handle.net/10616/45554

# Cognitive ability and risk for substance misuse in men: Genetic and environmental correlations in a longitudinal nationwide family study

Antti Latvala[1,2], Ralf Kuja-Halkola[2], Brian M D'Onofrio[3], Henrik Larsson[2], Paul Lichtenstein[2]

[1] Department of Public Health, PO Box 41, FI-00014 University of Helsinki, Finland

[2] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, PO Box 281, SE-17177 Stockholm, Sweden

[3] Department of Psychological and Brain Sciences, Indiana University- Bloomington, IN 47405, USA

CONTACT INFORMATION FOR CORRESPONDING AUTHOR (LATVALA):

Department of Public Health, PO Box 41, FI-00014 University of Helsinki

email: antti.latvala@helsinki.fi

Phone: +358 50 4484114

Fax: +358 2941 27518

Running head: Cognitive ability and substance misuse
Word count: 3500 / 3500
Competing interests: None

**Abstract**

**Aims.** To investigate the association between cognitive ability in late adolescence and subsequent substance misuse-related events in men, and to study the underlying genetic and environmental correlations. **Design.** A population-based longitudinal study with three different family-based designs. Cox proportional hazards models were conducted to investigate the association at the individual level. Bivariate quantitative genetic modeling in (1) full brothers and maternal half-brothers, (2) full brothers reared together and apart, and (3) monozygotic and dizygotic twin brothers was used to estimate genetic and environmental correlations. **Setting.** Register-based study in Sweden.

**Participants.** The full sample included 1,402,333 Swedish men born 1958-1991 and conscripted at mean age 18.2 (*SD*=0.5) years. 1,361,066 men who had no substance misuse events before cognitive assessment at mandatory military conscription were included in the Cox regression models with a follow-up time of up to 35.6 years. **Measures.** Cognitive ability was assessed at conscription with the Swedish Enlistment Battery. Substance misuse events included alcohol and drug related court convictions, medical treatments, and deaths, available from governmental registries **Findings.** Lower cognitive ability in late adolescence predicted an increased risk for substance misuse events (hazard ratio [HR] for a 1-stanine unit decrease in cognitive ability: 1.29, 95% CI: 1.29-1.30). The association was somewhat attenuated within clusters of full brothers (HR=1.21, 95% CI: 1.20-1.23). Quantitative genetic analyses indicated that the association was primarily due to genetic influences; the genetic correlations ranged between -.39 (95% CI: -.45, -.34) and -.52 (-.55, -.48) in the three different designs. **Conclusions.** Our findings from different family designs indicate that shared genetic influences underlie the association between low cognitive ability and subsequent risk for substance misuse events.

270 / 300 words

**Keywords**: Cognitive ability, substance misuse, quantitative genetic analysis, family study, longitudinal study, register-based research

## Introduction

Lower cognitive ability (CA) has been associated with an increased risk for mental disorders, including alcohol and drug use disorders [1-4]. Previous studies on CA and substance misuse have used psychiatric interviews [1] and registered medical information [2-4], and have assessed CA in childhood, adolescence or early adulthood. In these studies, CA has often been found to predict both alcohol and drug use disorders in a similar manner [2, 3]. Despite the associations being relatively well established, the underlying mechanisms have remained poorly understood. Familial factors are important for CA as well as alcohol and drug use, and twin and family studies indicate the familiality to be mainly due to genetic influences [5, 6]. Consequently, one possible cause for the association between low CA and risk for substance use disorders is pleiotropy, or shared genetic influences between the two traits [7].

Studies with twins have found support for shared genetic influences between CA and an alcohol use/problem composite measure [8], and between verbal CA and different indicators of alcohol problems [9]. However, in addition to genetic influences, environmental influences shared by family members contribute both to CA [6] and to drug abuse [10] in adolescence and young adulthood. Twin studies often have insufficient statistical power to detect small but non-negligible proportions of shared environmental variance [11], and previous studies may have been limited in their ability to detect shared environmental sources of covariance between CA and substance use problems [8, 9].

Further, despite the success of twin studies in estimating the heritability of human traits and behaviors [12], critics of the method have pointed out potential limitations, such as the equal environments assumption (EEA) [13] and representativeness of twin samples [14]. EEA is the assumption that etiologically relevant environments are no more similar in identical (monozygotic, MZ) than in fraternal (dizygotic, DZ) twins. Representativeness of twin studies

may be questioned on the basis of studies being typically based on voluntary participation which often results in substantial non-participation and attrition rates, and because twins may not be fully representative of the general population, although the latter critique seems invalid as twins have been shown to be similar to singletons in many psychological and medical traits [15-17].

Providing an accurate estimate of the risk for substance abuse associated with lower CA, as well as understanding why lower CA increases risk, are important research goals because of the currently limited understanding of individual differences in the risk for substance abuse. Specifically, understanding the causes of the familiality of the association is important, as genetic and shared environmental sources of covariance would imply different underlying risk mechanisms and point towards partly different intervention strategies. One approach to complement and extend previous twin studies is to use other family-based research designs, such as analyses of full and half-siblings [18-20].

To accomplish these goals, we conducted the largest study to date on the predictive association between CA and substance misuse-related events in medical and criminal registers, based on more than 35 years' coverage of nationwide register data from Sweden. We aimed (1) to describe the predictive association between CA in late adolescence and substance misuse events later in life, and (2) to clarify the genetic and environmental etiology of the association using quantitative genetic analyses based on three different family-based designs.

## Methods

*Study population*

We performed a nationwide cohort study of Swedish men by linking several longitudinal population registers maintained by governmental agencies. Data were available until December 31, 2009. A unique personal identification number, given to all citizens at birth,

was used as key in the register linkages. CA was assessed during a 2-day conscription assessment for the Swedish Armed Forces, mandatory until 2009 for all Swedish men at age 18. Only those with severe diseases, handicaps or intellectual disability were exempt, and more than 95% of men generally attended the conscription [21]. We linked information from the Total Population Register and the Conscription Register to identify all men who were born in Sweden between 1958 and 1991 and who were conscripted by end of 2009 (N = 1,488,886). The Multi-Generation Register (MGR) identifies biological and adoptive parents of each individual born since 1932 and living in Sweden at any time since 1961 [22]. MGR was used to link men with their biological and adoptive parents; those who had missing information on biological father's or mother's identity were excluded from the study (N = 16,413). Of the remaining 1,472,473 men 70,140 had a missing CA value. The study sample thus included 1,402,333 men.

In quantitative genetic analyses with three different family-based designs, data from relatives with at least 10 years of follow-up were used. Based on the MGR, 133,504 pairs of full brothers, 3163 pairs of maternal half-brothers, and 449 pairs of reared-apart full brothers were identified. From the Swedish Twin Registry [23], we identified 1,819 MZ and 1,700 DZ twin pairs with CA and substance misuse data. The study was approved by the institutional review board of the Karolinska Institutet.

*Measures*

Cognitive ability. CA was assessed with the Swedish Enlistment Battery (SEB). Three different versions of the SEB have been used during the 40-year period for which cognitive data are available: the SEB67 in 1970-1979, the SEB80 during 1980-1993, and the CAT-SEB for 1994-2009 [21]. The SEB67 and SEB80 were paper-and-pencil tests with four subtests assessing verbal, visuospatial, technical and inductive abilities, all summed to derive a general CA measure. High internal consistency has been reported for the SEB80 (coefficient alpha =

.79 - .91) [24]. A new version of the SEB (CAT-SEB), utilizing computer-aided testing, was launched in 1994. The CAT-SEB was based on a hierarchical model of cognitive abilities and included 12 tests, 10 of which were used to form a latent general CA factor. The reliability of the CAT-SEB tests is also good (coefficient alpha = .70 - .85) [25]. The CA variable, based on the described SEB versions, is presented on a stanine (standard nine) scale as a normally distributed variable divided into nine categories (1-9, lowest to highest cognitive performance) with a mean of 5.1 and a standard deviation of 1.9 in the present study. The scale was standardized separately for each year of conscripts, resulting in a constant distribution across the study period.

Substance misuse. Substance misuse events were assessed with an omnibus measure combining information about alcohol and drug related criminal convictions, medical treatments, and deaths. This measure, which has been used previously [10, 26, 27], captures a wide range of alcohol and drug related negative outcomes but does not include all individuals with substance abuse as many substance abusers are not registered in any medical or governmental registry. Court convictions of alcohol or drug related crimes included violations of the Narcotic Drugs Act (SFS 1968:64), as well as convictions of driving under the influence of alcohol and/or illicit substances, available from the National Crime Register for individuals aged 15 (age of criminal responsibility) and older since 1973. Data on medical treatments were available from the Patient Register, which contains details of all individual episodes of hospitalization in Sweden since 1973 and of outpatient treatments since 2001. Dates and causes of deaths were taken from the Cause of Death Register. Substance misuse events were defined based on the medical and death register data as having an International Classification of Diseases (ICD)-8/9/10 code related to alcohol or drugs, including mental and behavioral disorders due to alcohol and drugs (ICD-10: F10-9), poisonings (e.g. ICD-10: X45, X61-2, X65), as well as various somatic illnesses caused by alcohol or drug use (e.g. alcoholic

cardiomyopathy, alcoholic gastritis, and alcoholic liver disease [ICD-10: I42.6, K29.2, K70]). A complete list of the included ICD codes is given in supplementary Table S1.

Covariates. We included potential confounders of the association between CA and substance misuse. Socioeconomic status (SES) during childhood, available from National Censuses in 1960, 1970, 1980, 1985 and 1990, was derived from the occupation of the head of the household (usually the father) and coded into three classes: low (skilled and unskilled workers across all fields), medium (low- and intermediate-position white collar workers) and high (high-position white collar workers and self-employed professionals and entrepreneurs). Immigrant status for both parents was included as a dichotomous variable denoting whether the parent was born outside of Sweden. Linear effects of birth and conscription years were included to adjust for potential age and period effects.

*Statistical methods*

Individual-based analysis. We conducted Cox proportional hazard models to estimate the relative hazard of substance misuse events across the follow-up period starting from conscription. In addition to any substance misuse events, we conducted separate analyses for events based on medical and legal information to inspect differences between the two sources of information. Using medical data from the Patient and the Cause of Death registers, we also conducted separate models for alcohol- and drug-related events. The participants were followed until the occurrence of the first substance misuse event. Those who had no events within the study period contributed person-time at risk until the end of follow-up (December 31, 2009), emigration, or death, whichever occurred first. To reduce the possible effects of reverse causation (i.e., substance misuse affecting CA) and missing register data, we only included men who had no registered substance misuse events before conscription and who had not emigrated before conscription (N = 1,361,066).

The first Cox model adjusted for birth and conscription years, and the second model added childhood SES and parents' immigrant status. Graphical inspection of the Schoenfeld residuals for CA did not reveal violations of the proportional hazards assumption. The models were conducted with adjustment of standard errors for the non-independence of brothers using a robust sandwich estimator. For an indication of familial confounding, we conducted a stratified Cox regression within 271,113 clusters of full brothers (N = 570,963).

Quantitative genetic modeling. To estimate the contributions of genetic and environmental influences on the association between CA and substance misuse events, we conducted maximum-likelihood quantitative genetic structural equation modeling using three different family-based designs. This model decomposes the variance and covariance of the two traits into parts explained by additive genetic (*A*), shared environmental (*C*) and unique (i.e., individual-specific) environmental (*E*) influences [28]. Estimates from the model can be summarized as proportions of variance and covariance explained by the latent *A*, *C*, and *E* factors, and as correlations between the factors. Thus, a nonzero genetic correlation ($r_A$) between CA and substance misuse, for example, would imply that part of the genetic variation influencing CA also has an influence on the risk for substance misuse events. Similarly, correlations ($r_C$, $r_E$) for the two environmental variance components, *C* and *E*, would imply common environmental etiology between CA and substance misuse events, either shared by relatives (*C*) or unique to an individual (*E*).

The classical quantitative genetic model for twin data is defined by MZ and DZ co-twins having the correlations of 1 and .5, respectively, for the latent *A* factors due to the fact that MZ co-twins share 100% of their genome whereas DZ co-twins share, on average, 50% of their segregating genes [28]. Both MZ and DZ co-twins are set to have a correlation of 1 for the *C* factors, based on the assumption that they have been reared together. The *E* factors,

denoting individual-specific environmental influences (including measurement error), are defined to be uncorrelated for both twin types.

In addition to the model in twins (*Twin model*), we conducted corresponding bivariate models using data from (1) pairs of full brothers and maternal half-brothers (*Full/half model*), and (2) pairs of full brothers reared together and apart (*Full/adopted model*). In the *Full/half model*, *A* correlations for full brothers and maternal half-brothers were set to .5 and .25, respectively, and the *C* correlation was set to 1 for both groups. Thus, we assumed maternal half-brothers to have been reared together. This was based on statistics showing that in Sweden nearly all children have traditionally stayed with their mother after the parents' separation [29]. To make the assumption even more legitimate, we restricted the model to pairs in which brothers were born within less than 5 years of each other. For the *Full/adopted model*, we used information from the MGR to identify pairs of full brothers in which one brother had been adopted away and the other had not, or in which both brothers had been adopted away but into different homes. We assumed an *A* correlation of .5 for both groups, whereas *C* correlations were set to 1 and 0 for brothers reared together and apart, respectively. In all designs, *E* factors were defined as uncorrelated between brothers.

The CA variable was treated as continuous in the quantitative genetic models. For substance misuse events, a liability-threshold model was estimated using a binary variable which indicated substance misuse events within 10 years after conscription. Only men with at least 10 years of follow-up time were included in order to rule out bias arising from different length of follow-up. The liability-threshold model assumes each individual to have an unobserved normally distributed liability for substance misuse, and those with observed misuse events are assumed to have a liability exceeding the threshold [30].

Comparing with the full bivariate *ACE* model, we sequentially tested whether the *A* and *C* components for CA and substance misuse could be set to zero without statistically significant

deterioration in model fit, as indicated by likelihood ratio tests. Subsequently, we tested whether correlations between the statistically significant variance components (e.g. $r_A$) could be set to zero. The Akaike Information Criterion (AIC) was also used to assess model fit, with lower AIC values indicating better fit. We used the OpenMx package [31] in the software R [32] for model fitting.

Sensitivity analyses. We conducted separate individual-based analyses for the cohorts tested with the three versions of the SEB as well as for the four different cognitive subtests in SEB67 and SEB80. We repeated the main analysis including also substance misuse events that had occurred before conscription.

**Results**

Descriptive statistics by CA stanine and in the full sample are given in Table 1. The length of follow-up was on average 16.9 (SD=9.3) years with a maximum of 35.6 years. The rate of substance misuse events increased by decreasing CA, which is also indicated by the Kaplan-Meier survival curves (Figure 1).

[Table 1]

[Figure 1]

Adjusting for birth and conscription years, a 1-stanine-unit decrease in CA was associated with a 29% increase in the hazard for substance misuse in the Cox regression (Table 2). Adjustment for childhood SES and parental immigration had little effect, but the association was reduced within clusters of full brothers (Table 2), which suggests the presence of some familial confounding. The association was strongest for substance misuse events derived from the medical registers, and stronger for drug-related as compared to alcohol-related events.

[Table 2]

Correlations within pairs of brothers in all three designs suggested genetic influences on CA and substance misuse as well as their association (Tables S2-S4). Model fit and

comparisons of the bivariate quantitative genetic models are shown in Table 3. In all designs, significant *A* influences for both traits and a *C* component for CA were found, whereas *C* influences on substance misuse could be set to zero without significant deterioration in model fit. Thus, an $ACE_{CA}$-$AE_{substance\ misuse}$ model was selected for testing genetic and unique environmental correlations. In all designs, a statistically significant $r_A$ was present. A statistically significant $r_E$ was found in the *Full/half* and *Full/adopted* models but not in twins (Table 3). Proportions of variance explained by the *A*, *C*, and *E* factors, and the genetic and environmental correlations between CA and substance misuse from the best fitting models in the three designs are summarized in Figure 2 and in Table S5. The genetic correlation ranged between -.39 and -.52 in the three models. Estimates from the full bivariate models are given in Table S6.

[Table 3]

[Figure 2]

*Sensitivity analyses*

In sensitivity analyses, the association between CA and substance misuse was found to be similar for the paper-and-pencil (HR=1.29, 95% CI: 1.28-1.29) and computerized (HR= 1.31, 95% CI: 1.31-1.32) versions of the SEB.

Separate analyses of the four cognitive subtests in SEB67 and SEB80 data suggested a slightly stronger association for inductive and verbal domains than for visuospatial and technical domains (Table S7).

Including also substance misuse events that had occurred before conscription had no effect on the results (Table S8).

Finally, to assess the comparability of CA-substance misuse associations in the individual-based and quantitative genetic analyses, we conducted a probit regression analysis using the dichotomous substance misuse variable within 10 years of conscription. This resulted in a

similar association as the Cox regression (risk ratio [RR] associated with a 1-unit difference around the CA mean: RR=1.31, 95% CI: 1.31-1.32).

**Discussion**

Our large longitudinal cohort study provided unambiguous evidence for lower CA in late adolescence as a predictor of subsequent alcohol and drug misuse events. The observed hazard ratio for any substance misuse events corresponds to 62% increased risk per one standard deviation decrease in CA. This result, being based on almost 1.4 million individuals with follow-up up to 35 years, replicates and extends previous epidemiological studies on CA and substance abuse [1-4]. The association was stronger for medical events than for substance-related convictions, and it was stronger for drug-related than for alcohol-related medical events. Both findings are consistent with low CA being more strongly predictive of more severe substance misuse.

Furthermore, we found clear support for a negative genetic correlation between CA and risk for substance misuse events. Importantly, our analysis with three different family-based designs provided compatible estimates, with genetic correlations between -.4 and -.5. Thus, using both twin and sibling models we replicated results from earlier twin studies [8, 9].

Despite our large sample and different family designs, we did not find statistically significant shared environmental covariance between CA and substance misuse events. This was irrespective of shared environmental influences explaining 12-20% of variance in CA in late adolescence, which is in line with previous estimates [6, 33]. In contrast, there was no statistically significant shared environmental variance for substance misuse events in any of the three family-based designs. Our results suggest that the familiality of the association between lower CA and risk for substance misuse is due to overlapping genetic influences. The findings of a genetic correlation and limited evidence for shared environmental correlation are congruent with the previous twin studies [8, 9].

A recent study found support for genetic correlations between alcohol dependence and cognitive measures using polygenic risk scores based on genome-wide association studies of alcohol dependence [34]. This methodology is uninformative about environmental sources of the association but, using measured variation in the DNA, it confirms that the genetic correlations observed here and in previous studies are not artefacts produced by assumptions of quantitative genetic analyses of relatives.

Interestingly, we found a statistically significant positive non-shared environmental correlation between CA and substance misuse events in two of the three designs. Such positive correlation has been reported earlier [8]. The finding implies that in addition to shared genetic variation which makes it more likely that an individual has both lower CA and higher risk for substance misuse-related events, there are counterbalancing environmental influences which are specific to each individual. However, the negative genetic association outweighs this effect and leads to an overall negative phenotypic association.

Limitations of our study include that the sample was entirely comprised of men; the results may not be fully generalizable to women. Second, we used register-based data on alcohol and drug misuse, which may have resulted in exclusion of less severe cases and led to biased estimates. However, we combined data from medical and criminal registers in order to capture more instances of serious substance misuse-related events. Third, our data did not include information about levels of alcohol or drug use before or after conscription. However, when we included register-based substance misuse events that had occurred before the assessment of CA as events in the Cox model, the results were unchanged. This, in combination with the known high stability of CA [35], suggests that our results reflect a true prospective association between CA and later substance misuse events. Further, although our analyses indicated shared genetic influences we cannot rule out a causal association. Finally, all statistical designs have assumptions and limitations, including the three quantitative genetic designs that

were used in this report. However, because the three designs are based on different assumptions, it is unlikely that shared environmental influences explain a substantial part of the associations between CA and substance misuse events.

In conclusion, our findings provide further support for a prospective association between lower CA and an increased risk for substance misuse events. We found that familial factors contributed to this association, and quantitative genetic analyses using three different designs indicated that the familiality was primarily due to correlated genetic influences on CA and substance misuse. Lower CA should be considered as an additional marker of risk for alcohol and drug use problems, and unfolding of the association should be studied further in developmental settings and with genetically informative data.

**Acknowledgements.**

This study was supported by grants from the Academy of Finland, the Swedish Research Council for Health, Working Life and Welfare (2012-1678), the Swedish Research Council (2011-2492), Swedish Research Council through the Swedish Initiative for Research on Microdata in the Social And Medical Sciences (SIMSAM) (340-2013-5867), and the European Union Seventh Framework Programme (FP7/2007-2013; 602768). The sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

## References


1. Gale CR, Deary IJ, Boyle SH, Barefoot J, Mortensen LH, Batty GD. Cognitive ability in early adulthood and risk of 5 specific psychiatric disorders in middle age. *Archives of General Psychiatry* 2008; 65: 1410-1418.
2. Gale CR, Batty GD, Tynelius P, Deary IJ, Rasmussen F. Intelligence in early adulthood and subsequent hospitalization for mental disorders. *Epidemiology* 2010; 21: 70-77.
3. Mortensen EL, Sorensen HJ, Jensen HH, Reinsisch JM, Mednick SA. IQ and mental disorders in young men. *British Journal of Psychiatry* 2005; 187: 407-415.
4. Sjölund S, Allebeck P, Hemmingson T. Intelligence quotient (IQ) in adolescence and later risk of alcohol-related hospital admissions and deaths—37-year follow-up of Swedish conscripts. *Addiction* 2011; 107: 89-97.
5. Agrawal A, Verweij KJH, Gillespie NA, Heath AC, Lessov-Schlaggar CN, Martin NG, et al. The genetics of addiction—a translational perspective. *Translational Psychiatry* 2012; 2: e140.
6. Haworth CM, Wright MJ, Luciano M, Martin NG, de Geus EJ, van Beijsterveldt CE, et al. The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry* 2010; 15: 1112-1120.
7. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. *Nature Reviews Genetics* 2013; 14: 483-495.
8. Johnson W, Hicks BM, McGue M, Iacono WG. How intelligence and education contribute to substance use: Hints from the Minnesota Twin Family Study. *Intelligence* 2009; 37: 613-624.

9. Latvala A, Tuulio-Henriksson A, Dick DM, Vuoksimaa E, Viken RJ, Suvisaari J, et al. Genetic origins of the association between verbal ability and alcohol dependence symptoms in young adulthood. *Psychological Medicine* 2011; 41: 641-651.
10. Kendler KS, Maes HH, Sudquist K, Olsson H, Sundquist J. Genetic and family and community environmental effects on drug abuse in adolescence: a Swedish national twin and sibling study. *American Journal of Psychiatry* 2014; 171: 209-217.
11. Visscher PM, Gordon S, Neale MC. Power of the classical twin design revisited: II detection of common environmental variance. *Twin Research and Human Genetics*; 2008 11: 48-54.
12. Polderman TJC, Benyamin B, de Leeuw CA, Sullivan PF, van Bochoven A, Visscher PM, Posthuma D. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics* 2015; 47: 702-709.
13. Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. A test of the equal-environment assumption in twin studies of psychiatris illness. *Behavior Genetics* 1993; 23: 21-27.
14. Evans DM, Martin NG. The validity of twin studies. *GeneScreen* 2000; 1: 77-79.
15. Andrew T, Hart DJ, Snieder H, de Lange M, Spector TD, MacGregor AJ. Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. *Twin Research* 2001; 4: 464-477.
16. Christensen K, Vaupel JW, Holm NV, Yashin AI. Mortality among twins after age 6: Fetal origins hypothesis versus twin method. *BMJ* 1995; 310: 432-436.
17. Johnson W, Krueger RF, Bouchard TJ, McGue M. The personalities of twins: Just ordinary folks. *Twin Research* 2002; 5: 125-131.

18. D'Onofrio BM, Lahey BB, Turkheimer E, Lichtenstein P. Critical need for family-based, quasi-experimental designs in integrating genetic and social science research. *American Journal of Public Health* 2013; 103: S46-S55.
19. Falconer DS, Mackay TFC. *Introduction to Quantitative Genetics, Fourth Edition.* Harlow, England: Pearson Education Limited 1996.
20. Rutter M. Proceeding from observed correlation to causal inference: the use of natural experiments. *Perspectives on Psychological Science.* 2007; 2: 277-395.
21. Carlstedt B. *Cognitive abilities - aspects of structure, process and measurement*. Göteborg: Göteborgs Universitet 2000.
22. Statistics Sweden. *Multi-generation register 2009. A description of contents and quality.* Örebro, Sweden: Statistics Sweden 2010.
23. Lichtenstein P, Sullivan PF, Cnattingius S, et al. The Swedish Twin Registry in the third millennium: an update. *Twin Research and Human Genetics*. 2006; 9:875-882.
24. Carlstedt B, Mårdberg B. Construct validity of the Swedish Enlistment Battery. *Scandinavian Journal of Psychology* 1993; 34: 353–362.
25. Mårdberg B, Carlstedt B. Swedish Enlistment Battery (SEB): Construct validity and latent variable estimation of cognitive abilities by the CAT-SEB. *International Journal of Selection and Assessment* 1998; 6: 107–114.
26. D'Onofrio BM, Rickert ME, Långström N, Donahue KL, Coyne CA, Larsson H, et al. Familial confounding of the association between maternal smoking during pregnancy and offspring substance use and problems. *Archives of General Psychiatry* 2012; 69: 1140-1150.
27. Sariaslan A, Långström N, D'Onofrio BM, Hallqvist J, Franck J, Lichtenstein P. The impact of neighbourhood deprivation on adolescent violent criminality and substance

misuse: A longitudinal, quasi-experimental study of the total Swedish population. *International Journal of Epidemiology*. 2013; 42: 1057-1066.

28. Neale MC, Cardon LR. *Methodology for genetic studies of twins and families*. Dordrecht, The Netherlands: Kluwer Academic Publishers 1992.
29. Statistics Sweden. *Fakta om den svenska familjen*. SCB Demografiska rapporter (Facts about the Swedish Family. SCB Demograpghic Reports). Statistics Sweden 1994.
30. Posthuma D, Beem AL, de Geus EJC, van Baal GCM, von Hjelmborg JB, Iachine I, Boomsma DI. Theory and practice in quantitative genetics. *Twin Research.* 2003; 6: 361-376.
31. Boker SM, Neale MC, Maes HH, Wilde MJ, Spiegel M, Brick TR, et al. *OpenMx 1.2.0 user guide*. 2012. Retrieved from http://openmx.psyc.virginia.edu/documentation
32. R Development Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing 2012.
33. Shakeshaft NG, Trzaskowski M, McMillan A, Krapohl E, Simpson MA, Reichenberg A, et al. Thinking positively: The genetics of high intelligence. *Intelligence* 2015; 48: 123-132.
34. Clarke TK, Smith AH, Gelernter J, Kranzler HR, Farrer LA, Hall LS, et al. Polygenic risk for alcohol dependence associates with alcohol consumption, cognitive function and social deprivation in a population-based cohort. *Addiction Biology* 2015; Epub ahead of print.
35. Deary IJ. Intelligence. *Annual Review of Psychology*. 2012; 63: 453–482.

**Figure 1.** Kaplan-Meier survival curves for substance misuse by cognitive ability (CA) stanines

**Figure 2.** Standardized estimates (95% CIs) from the best fitting bivariate quantitative genetic models of cognitive ability (CA) and substance misuse (SM) in three family-based designs

*Notes for Figure 2: A = additive genetic variance, C = shared environmental variance, E = unique (individual-specific) environmental variance, $r_A$ = additive genetic correlation, $r_E$ = unique environmental correlation. Confidence intervals are of Wald type hence they may expand outside the parameter space, standard errors are calculated using the delta method.*

**Table 1.** Rate of substance misuse events by cognitive ability in Swedish men

| Cognitive ability score (stanine) | No of men | Mean (SD) age at conscription (years) | Person-years at risk | No of substance misuse events | Rate (95% CI) per 10,000 person-years at risk |
|---|---|---|---|---|---|
| 1 | 40,683 | 18.2 (0.6) | 616,337 | 8,059 | 131 (128–134) |
| 2 | 88,865 | 18.2 (0.5) | 1,448,034 | 14,694 | 101 (100–103) |
| 3 | 147,319 | 18.2 (0.5) | 2,365,315 | 19,787 | 84 (82–85) |
| 4 | 213,626 | 18.2 (0.5) | 3,489,010 | 22,965 | 66 (65–67) |
| 5 | 304,287 | 18.2 (0.4) | 5,246,079 | 25,372 | 48 (48–49) |
| 6 | 233,899 | 18.2 (0.4) | 4,007,343 | 14,766 | 37 (36–37) |
| 7 | 172,594 | 18.2 (0.4) | 3,003,448 | 8,284 | 28 (27–28) |
| 8 | 103,157 | 18.2 (0.4) | 1,817,107 | 3,841 | 21 (20–22) |
| 9 | 56,636 | 18.3 (0.5) | 997,164 | 1,525 | 15 (15–16) |
| Total sample (*M*=5.1, *SD*=1.9) | 1,361,066 | 18.2 (0.5) | 22,989,840 | 119,293 | 52 (52–52) |

Note: Cognitive ability stanines range from 1 (lowest ability group) to 9 (highest ability group)

**Table 2.** Cox proportional hazard ratios (95% confidence intervals) for substance misuse events by 1 stanine-unit decrease in cognitive ability

| | Model adjusted for birth and conscription years | Model additionally adjusted for childhood SES and parents' immigrant status | Stratified model within clusters of full brothers, adjusted for birth and conscription years |
|---|---|---|---|
| Any substance misuse event | 1.29 (1.29–1.30) | 1.28 (1.27–1.28) | 1.21 (1.20–1.23) |
| Substance misuse event based on medical registers | 1.34 (1.33–1.35) | 1.33 (1.32–1.33) | 1.25 (1.23–1.27) |
| *Alcohol misuse events* | 1.32 (1.31-1.32) | 1.30 (1.29-1.31) | 1.23 (1.21-1.25) |
| *Drug misuse events* | 1.46 (1.45-1.48) | 1.44 (1.43-1.46) | 1.36 (1.32-1.39) |
| Substance misuse event based on court convictions | 1.32 (1.31–1.31) | 1.29 (1.29–1.30) | 1.22 (1.20–1.23) |

SES, socioeconomic status

**Table 3.** Bivariate quantitative genetic models of cognitive ability (CA) and substance misuse (SM) in three family-based designs

| Relative groups and model | *df* | AIC | -2LL | Comparison model | -2LL difference | *p* |
|---|---|---|---|---|---|---|
| Full brothers and maternal half-brothers | | | | | | |
| 1. Full bivariate model | 715,881 | 318,907.2 | 1,750,669.2 | - | - | - |
| 2. $A_{CA} = 0$ | 715,883 | 319,091.2 | 1,750,857.2 | 1 | 188.03 | <0.001 |
| 3. $A_{SM} = 0$ | 715,883 | 318,933.3 | 1,750,699.3 | 1 | 30.17 | <0.001 |
| 4. $C_{CA} = 0$ | 715,883 | 318,930.9 | 1,750,696.6 | 1 | 27.47 | <0.001 |
| **5. $C_{SM} = 0$** | **715,883** | **318,903.8** | **1,750,669.8** | **1** | **0.68** | **0.710** |
| 6. $r_A = 0$ | 715,884 | 320,952.1 | 1,752,720.1 | 5 | 2,050.9 | <0.001 |
| 7. $r_E = 0$ | 715,884 | 319,098.1 | 1,750,866.1 | 5 | 196.91 | <0.001 |
| Full brothers reared together and apart | | | | | | |
| 1. Full bivariate model | 697,033 | 307,623.3 | 1,701,689.3 | - | - | - |
| 2. $A_{CA} = 0$ | 697,035 | 307,682.0 | 1,701,752.0 | 1 | 62.70 | <0.001 |
| 3. $A_{SM} = 0$ | 697,035 | 307,628.3 | 1,701,698.3 | 1 | 8.95 | 0.011 |
| 4. $C_{CA} = 0$ | 697,035 | 307,627.6 | 1,701,697.6 | 1 | 8.25 | 0.016 |
| **5. $C_{SM} = 0$** | **697,035** | **307,620.2** | **1,701,690.0** | **1** | **0.74** | **0.692** |
| 6. $r_A = 0$ | 697,036 | 309,630.0 | 1,703,702.0 | 5 | 2,012.0 | <0.001 |
| 7. $r_E = 0$ | 697,036 | 307,800.3 | 1,701,872.3 | 5 | 174.75 | <0.001 |
| MZ and DZ twin brothers | | | | | | |
| 1. Full bivariate model | 17,112 | 4,606.9 | 38,830.9 | - | - | - |
| 2. $A_{CA} = 0$ | 17,114 | 5,073.2 | 39,301.2 | 1 | 470.32 | <0.001 |
| 3. $A_{SM} = 0$ | 17,114 | 4,611.9 | 38,839.9 | 1 | 9.08 | 0.011 |
| 4. $C_{CA} = 0$ | 17,114 | 4,643.9 | 38,871.9 | 1 | 41.03 | <0.001 |
| 5. $C_{SM} = 0$ | 17,114 | 4,607.2 | 38,835.2 | 1 | 4.37 | 0.112 |
| 6. $r_A = 0$ | 17,115 | 4,665.8 | 38,895.8 | 5 | 60.55 | <0.001 |
| **7. $r_E = 0$** | **17,115** | **4,605.6** | **38,835.6** | **5** | **0.38** | **0.538** |

CA = cognitive ability, SM = substance misuse, df = degrees of freedom, AIC = Akaike information criterion, -2LL = -2 log-likelihood, *A* = additive genetic variance, *C* = shared environmental variance, *E* = unique (individual-specific) environmental variance, $r_A$ = additive genetic correlation, $r_E$ = unique environmental correlation
The best fitting model in each design is shown in bold font.