Published in final edited form as:

Addiction. 2013 October ; 108(10): 1767–1776. doi:10.1111/add.12187.

Measurement invariance of DSM-IV alcohol, marijuana, and cocaine dependence between community-sampled and clinically over-selected studies

Jaime Derringer^{1,*}, Robert F. Krueger², Danielle M. Dick³, Arpana Agrawal⁴, Kathleen K. Bucholz⁴, Tatiana Foroud⁵, Richard A. Grucza⁴, Michie N. Hesselbrock⁶, Victor Hesselbrock⁶, John Kramer⁷, John I. Nurnberger Jr.⁵, Marc Schuckit⁸, Laura J. Bierut⁴, William G. Iacono², and Matt McGue²

¹University of Colorado

²University of Minnesota

³Virginia Commonwealth University

⁴Washington University in St. Louis

⁵Indiana University

⁶University of Connecticut

⁷University of Iowa

⁸University of California San Diego

Abstract

Aims—To examine whether DSM-IV symptoms of substance dependence are psychometrically equivalent between existing community-sampled and clinically over-selected studies.

Participants—2476 adult twins born in Minnesota; 4121 unrelated adult participants from a case-control study of alcohol dependence.

Measurements—Lifetime DSM-IV alcohol, marijuana, and cocaine dependence symptoms and ever use of each substance.

Design—We fit a hierarchical model to the data, in which ever use and dependence symptoms for each substance were indicators of alcohol, marijuana, or cocaine dependence, which were in turn indicators of a multi-substance dependence factor. We then tested the model for measurement invariance across participant groups, defined by study source and participant sex.

Findings—The hierarchical model fit well among males and females within each sample (CFI>0.96, TLI>0.95, and RMSEA<0.04 for all), and a multi-group model demonstrated that model parameters were equivalent across sample- and sex-defined groups (CFI=0.002 between constrained and unconstrained models). Differences between groups in symptom endorsement rates could be expressed solely as mean differences in the multi-substance dependence factor.

Conclusions—Lifetime substance dependence symptoms fit a dimensional model well. Although clinically over-selected participants on average endorsed more dependence symptoms than community-sampled participants, the pattern of symptom endorsement was similar across

* Corresponding author: Jaime Derringer, Institute for Behavioral Genetics, University of Colorado, Boulder, CO 80309-0447; jaime.derringer@colorado.edu.

groups. From a measurement perspective, DSM-IV criteria are equally appropriate for describing substance dependence across different sampling methods.

Keywords

substance dependence; sampling comparison; sex differences; item response theory

Measurement invariance of DSM-IV alcohol, marijuana, and cocaine dependence between community-sampled and clinically over-selected studies

The structure of substance use and substance use disorders (SUDs - including abuse and dependence), as defined by the Diagnostic and Statistical Manuals (DSMs) [1], both for individual and multiple substances, has been modeled in a range of populations using a variety of measurement techniques. Here we aim to test whether the structure of substance dependence criteria endorsement across multiple substances is similar between different sampling methods (i.e., community-sampled twins versus clinically over-selected individuals) as well as between sexes. Applying diagnostic criteria across a variety of research settings assumes implicitly that these criteria have equivalent measurement properties (such as rank order of symptom endorsement frequencies) when applied to either clinical or epidemiological samples. This is not a minor issue, given that DSM criteria are explicitly designed for application within clinical rather than epidemiological settings, but often studied in general population and research-based samples.

Many previous studies identified a unidimensional structure of SUDs for individual substances. These findings are robust enough that a unidimensional structure of SUDs will be implemented in the DSM-5. [2] DSM-5 eliminates the abuse/dependence distinction and describes SUDs as a continuum, by including severity specifiers such as moderate versus severe, rather than just a categorical diagnosis. [3,4] Previous research strongly supports a unidimensional structure of DSM-IV SUD symptoms within specific substances, including alcohol [5–9], marijuana [8–11], and cocaine [9–11]. Non-DSM measures of problems with alcohol [12–16] and marijuana [17–18] also show evidence of unidimensionality. While many of these studies relied on self-report inventories, unidimensionality of substance use and problems has also been reported for a combination of self-report problem inventories with confirmation of use by biologic markers of substance use (e.g., saliva or urine samples) for marijuana, cocaine, and other illicit substances. [19] This wealth of prior research supports the unidimensionality of substance use and problems within individual substances. We turn now to considering models of multi-substance use, as well as how these models compare across different demographic groups and sampling schemes.

Structure of multiple SUDs

SUD symptoms for individual substances are consistently found to be unidimensional. Further, problems with multiple substances may also represent a single continuum. SUDs tend to positively correlate across a wide range of specific substances. That is, problems with any given substance often predict a common liability to SUDs in general, rather than liability to problems with only a single, specific substance. Certain risk factors contribute to risk of problems with a range of substances. For example, early-adolescent characteristics (such as aggression and delinquency) predict increased rates of SUDs in early adulthood, but do not differentiate between substances. That is, we may expect that a person exposed to these risk factors is at an increased risk of any SUD, but we have limited ability to predict which specific substance(s) will become problematic. [20]

The consistently high correlations among SUDs and presence of at least some common risk factors suggests that multiple substances may form a single underlying multi-substance dependence continuum. In a study by Kirisci and colleagues [21], multiple SUD diagnoses fit a unidimensional model in adult men (over-sampled for SUDs), their wives, and their adolescent sons. A single dimension also captures ever use of multiple substances, for both males and females in adolescent [22] and adult samples. [23] These studies demonstrate that the unidimensional structure of multiple substance problems is similar across ages and sexes, although they do not address concerns of generalizing the structure of substance dependence across samples drawn from different populations.

Studies comparing sampling methods

While previous studies examined the issues of age and sex differences (or similarities) in the factor structure of SUDs, there have been far fewer investigations comparing the factor structure of SUDs across multiple sampling schemes. Studies comparing the structure of non-substance characteristics between population and clinical samples found that the same factor structure applies when examining measures of intelligence [24], broad neuropsychological batteries [25], and death distress. [26] Here, clinic-based and population samples are not distinguishable by their measurement structures, but rather simply differ in their average trait levels. Conversely, a measure of alexithymia (difficulty understanding emotions) displays a slightly different factor structure when comparing clinical and volunteer samples. [27]

Few studies examined how the factor structure of substance use or related problems generalizes across differing sampling schemes. In a comparison of substance use behaviors between volunteer and random samples, volunteers reported a greater incidence of psychosocial risk factors (such as lower average socioeconomic status and IQ) and higher rates of SUDs for “hard” drugs, but the two groups were not distinguishable on the basis of overall number or severity of SUDs. Further, volunteers demonstrated less social desirability bias in self-reports. [28] Comparing drug use problems between Swedish heavy drug users and a population sample, a drug problems inventory best fit a three-factor solution among the heavy drug using sample (including substance treatment in-patients and individuals who were either incarcerated or on probation), but a two-factor solution in the population sample. However, the authors of this particular study noted that the low prevalence of substance-related problems in the population sample suggest that factor results may be unreliable in that sample. [29]

The current study

In the current analyses, we examine three questions relevant to the conceptualization of substance dependence criteria across multiple substances, in the context of comparing different substances across sexes and type of sample. First, we test whether a hierarchical model adequately describes substance dependence criteria for alcohol, marijuana, and cocaine, where substance dependence symptoms (as well as ever use of each substance) are indicators of individual substance factors, which in turn are indicators of a multi-substance dependence factor. Second, we examine whether this model demonstrates similar model-fit properties when comparing males and females in two independent studies, designed as either community-sampled or clinically over-selected for substance dependence. Finally, we test whether differences between samples can be solely attributed to mean differences in a multi-substance dependence latent trait rather than differing factor structures or substance-specific differences.

Applying a factor model to these data to test measurement invariance across sex- and sample-defined groups enables us to explicitly test the assumption that these DSM

dependence criteria apply equally well (from a psychometric perspective) to epidemiological and clinically over-selected samples. Because research often seeks to generalize findings between different populations, we must establish that DSM criteria measure the same constructs in the same way across different samples. This has not, to our knowledge, been established for DSM substance dependence criteria across clinical and community samples and is the primary aim of the current research.

Methods

Participants

MTFS—The Minnesota Twin Family Study (MTFS) [30,31] includes twins born in Minnesota and recruited to participate in longitudinal assessments. Two cohorts (a younger cohort who entered the study at age 11, and an older cohort who began at age 17) participated in repeated assessments approximately every 3–4 years around ages 11, 14, 17, 21, and 25. We identified all twins assessed for substance dependence symptoms through their 25-year-old follow-up visit. The MTFS sample ($N = 2476$) had a mean age of 25.00 ($SD = 0.89$, range = 22–29), was 52.5% female and primarily Caucasian, and was demographically representative of the Minnesota population. Through their age 25 assessment, 11.5% of MTFS participants met DSM-IV criteria for alcohol dependence at some point in their lifetime, 7.7% met criteria for marijuana dependence, and 1.6% met criteria for cocaine dependence.

SAGE—The Study of Addiction: Genes and Environment (SAGE) [32] is a case-control sample of unrelated individuals over-selected for substance dependence. Specifically, the SAGE sample was derived from three primary studies of alcohol, nicotine, and cocaine dependence, such that the final sample included approximately 50% alcohol dependent cases and 50% non-dependent controls. All controls were community-ascertained and did not meet criteria for dependence on any substance. Alcohol dependent cases drawn from the primary study of nicotine dependence were community-ascertained, while alcohol dependent cases drawn from the primary studies of alcohol and cocaine dependence were clinically ascertained. All alcohol dependent cases qualified for inclusion in SAGE regardless of dependence status on other substances. The SAGE study ($N=4121$) is 54.3% female, 67.3% Caucasian and 32.5% African American, with 3.4% participants reporting Hispanic ancestry. The SAGE sample has a mean age of 39.03 ($SD = 9.10$, range = 18–77), with 47.2% participants meeting DSM-IV criteria for alcohol dependence, 18.3% meeting criteria for marijuana dependence, and 27.4% meeting criteria for cocaine dependence. All data from SAGE utilized for the current study are publicly available via the database of Genotypes and Phenotypes (dbGaP; phs000092.v1.p1).

Measures

We selected three substances (alcohol, marijuana, and cocaine) for analysis because they represent a range of availability, severity, and legal statuses, and both the MTFS and SAGE studies assessed these substances in the same manner. For each substance, both studies assessed ever use of the substance and endorsement of each of the seven DSM-IV dependence criteria (in SAGE, via the SSAGA-II [33,34]; and in the MTFS, via the CIDI-SAM [35]). For the MTFS, lifetime ever use and dependence symptom endorsement was established by aggregating across all longitudinal assessments through the age 25 assessment. That is, ever use or any substance dependence symptom was considered “present” if the participant met the criterion at any assessment. Although not pathological, we include the ever use criterion to differentiate individuals exposed to a substance who have no dependence symptoms from those who were never exposed. Substance abuse

symptoms were not included in the present analyses, as equivalent abuse criteria were not assessed in both the SAGE and MTFS samples.

Analyses

Based on substantial prior research indicating that 1) substance dependence symptoms are unidimensional within each substance, and 2) multiple substances are indicators of a single latent trait of multi-substance dependence, we fit a hierarchical confirmatory factor model (depicted in Figure 1). Within this model, substance-specific dependence symptoms (along with ever use of the substance) were indicators of a factor specific to their respective substance (i.e., alcohol, marijuana, or cocaine). The estimation of a continuous factor model of symptoms, rather than restriction to dichotomous diagnoses, makes full use of the available information. Those substance-specific factors were in turn indicators of a higher-order multi-substance dependence factor. In modeling dichotomous substance criteria as indicators of a continuous latent trait, we estimate threshold parameters, in addition to the loading parameters estimated for both categorical and continuous indicators. Thresholds represent the standardized latent trait level (Z-score) at which the probability of an individual endorsing that criterion is 50%; therefore, higher thresholds represent less frequently endorsed criteria.

We examined model fit in three scenarios. In the first, we modeled all data concurrently, regardless of group membership, where “group” is defined by sex (male versus female) and study source (MTFS, community-sampled; versus SAGE, clinically over-selected). Second, we fit the model separately in each group. From these initial models we evaluated whether the proposed hierarchical factor model fit the data appropriately within each of the sex- and sample-defined groups, prior to formally testing measurement invariance in a multigroup model.

Next, we examined a multigroup model, in which group differences were captured either by differences in criterion parameters (in the unconstrained model), or by mean differences in the multi-substance factor (in the constrained model). In the first multigroup model, the “unconstrained” model, thresholds (for the categorical observed criteria) and factor loadings were estimated freely within each group, while residual variances and factor means were fixed (at one and zero, respectively) in all groups. In the second multigroup model, the “constrained” model, thresholds (for the categorical observed criteria) and factor loadings were constrained equal across all groups, while residual variances and the multi-substance latent trait mean were fixed (at one and zero, respectively) only in the reference group (MTFS males) and freely estimated in all other groups. The comparison of these two models allowed us to evaluate measurement invariance between the different sex and sampling groups. Because only the means (and residual variances) vary among groups in the constrained model, it is more parsimonious, but assumes that the substance dependence criteria exhibit measurement invariance across groups. That is, the structure of the substance dependence “measure” is assumed to be identical across groups, and any group differences in symptom endorsement rates are captured simply by mean differences in the multi-substance dependence latent trait.

We estimated all models in Mplus [36] using a weighted least squares estimator with theta parameterization. We used clustering (based on family membership) to account for the non-independence of the twin observations within the MTFS sample. We considered absolute model fit in terms of RMSEA (with values less than 0.06 indicating a well-fitting model), as well as CFI and TLI (with values greater than 0.95 indicating a well-fitting model [37]). We evaluated relative model fit by the difference in CFI values. If the difference in CFI values between constrained and unconstrained models was less than 0.01, or if the fit indices were

substantially better in the simpler constrained model, we retain the null hypothesis that the factor structure is the same between groups. [38]

Results

Table 1 provides the endorsement rates of ever use and the seven dependence symptoms for alcohol, marijuana, and cocaine in subgroups defined by sex and sample (that is, MTFS males, MTFS females, SAGE males, and SAGE females). For both the MTFS and SAGE samples, males endorsed criteria a median of twice as frequently as females, and this pattern was consistent across substances. SAGE participants endorsed dependence symptoms a median of three times as frequently as MTFS participants (comparing across samples within sexes). Between-sample differences in symptom endorsement rates varied more greatly by substance (compared to the relative consistency across substances observed when examining endorsement rates by sex). In particular, cocaine symptoms showed the greatest difference between SAGE and MTFS samples, due in part to similar disparities in endorsement rates of having ever used cocaine.

Table 2 provides fit statistics for each model applied to the data (or subsets thereof) and correlations among all substance dependence criteria are provided in Supplemental Tables S1 and S2. The hierarchical model presented in Figure 1 fit well when modeling all data as a single sample. Similarly, model fit was good within each of the individual groups (although the ever use criterion for cocaine was dropped from the model for SAGE males, due to multicollinearity between this and other criteria that prevented the model from converging on a solution).

When applying the hierarchical model in a multigroup context, both the unconstrained model (in which criterion loadings and thresholds were estimated separately for each group) and the constrained model (in which group means for multi-substance dependence varied among groups, while holding constant the criterion loadings and thresholds) fit the data well, as indicated by low RMSEA and high CFI and TLI. The difference in CFI between the constrained and unconstrained models was 0.002 (which is less than the recommended maximum difference cut-off of 0.01 for establishing measurement invariance [38]). Therefore, we conclude that mean differences in the multi-substance dependence latent trait adequately describe criterion-level differences between the groups and proceed with interpreting the results of the constrained model.

Table 3 provides model parameters from the preferred constrained model. Figure 2 illustrates the criterion loadings and thresholds for each individual substance as Criterion Information Curves (see Embretson & Reise [39] for a thorough description of the Item Response Theory framework that defines these plots). Within Figure 2, criteria that relate more strongly to the substance-specific factor (i.e., with greater loadings) display higher peaks, while criteria typically endorsed by individuals with a more severe level of the substance-specific latent trait are located farther to the right (i.e., with higher thresholds, where the x-axis is a Z-score metric).

Figure 2 demonstrates visually that, as expected based on endorsement rates, ever use or endorsement of dependence symptoms for cocaine are more severe indicators of substance involvement (with ever use located approximately 2.0 standard deviations above the mean, and dependence symptoms located between 3.0 and 4.0 standard deviations above the mean), compared to either alcohol or marijuana (ever use of either of which is common, even among the community-sampled MTFS males and females). Similarly, as alcohol use and endorsement of dependence symptoms is far more common than endorsement of similar criteria for marijuana across all groups, these plots illustrate that individuals endorsing the

same criteria represent a less extreme (that is, more normative) level of alcohol involvement (for which dependence symptoms are located between 0.5 and 2.5 standard deviations above the mean) compared to individuals endorsing similar criteria for marijuana (for which dependence symptoms are located between 2.5 and 3.5 standard deviations above the mean). These patterns reflect the “dependence liability” of each substance (e.g., alcohol dependence is “easier” to achieve than marijuana dependence). They do not capture differences in harm or treatment potential that may also vary between substances.

Within the preferred constrained model, mean differences in the multi-substance factor entirely capture the differences in individual substance criteria endorsement between groups. The mean multi-substance latent trait level was -0.57 for MTFS females, 0.00 for MTFS males (fixed within the model as the reference group), 0.51 for SAGE females, and 1.30 for SAGE males. These mean differences in the multi-substance factor mirror group differences in criteria endorsement rates (which were consistent across substances, see Table 1). The standardized loadings of all three substance-specific factors on the multi-substance factor were high (ranging from 0.72 to 0.89), indicating that the multi-substance factor accounts for a majority of the variance in ever use and dependence symptom endorsement rates.

Discussion

Based on evidence from the existing substance dependence literature, we identified a hierarchical factor model that provided good fit to alcohol, marijuana, and cocaine dependence symptom data in both a community-sampled and a clinically over-selected study. Measurement invariance of the model across samples and between sexes demonstrates that patterns of differences in criterion endorsement rates between these groups are primarily explained by mean differences in a higher-order multi-substance dependence trait, rather than being specific to individual substances or symptom parameters. Our findings support the psychometric validity of combining population and clinical samples in explorations of substance dependence etiology and outcomes.

Limitations

These findings should be considered in light of several limitations, which may impact the generalizability of our conclusions. Neither of these samples were representative of the U.S. population as a whole, and our model does not include the effects of either ethnicity or age, both of which were more diversely represented in our clinically over-selected sample. Previous studies report similar factor structures across ages [21], as well as among African-, European-, and Mexican-American individuals for alcohol [40] and cocaine dependence. [41] Our finding of measurement invariance in substance dependence across samples in the current study, despite these additional potential confounds not being explicitly modeled, strengthens our conclusion, based on this and previous studies, that the factor structure of substance dependence is consistent across samples with varying demographic characteristics.

We also note that the current findings refer to lifetime history of substance use and dependence symptoms and so may not generalize to models of recent substance problems, where co-morbidity among substances or even symptoms within a substance may be reduced. Although we include the ever use criterion as a broad indicator of whether or not an individual has tried a substance, there remains a gap in information available in the current study between ever use of a substance and the development of dependence symptoms. Further, there are likely additional key aspects of addiction, such as treatment-relevant factors like social support, that the DSM-IV dependence criteria do not capture.

Conclusions

The current study provides strong support for the conceptualization of substance dependence as a dimensional construct, both in terms of individual substances and at a higher-order level encompassing a range of multiple substances. Factor loadings indicate that the higher-order multi-substance dependence factor explained 50–80% of the observed variance in each of the substance-specific dependence factors. The construct of multi-substance dependence has been previously and consistently shown to be an indicator of the broader externalizing spectrum, encompassing a range of substance use and misuse measures, non-substance disinhibitory behaviors (including criteria for Conduct Disorder and Antisocial Personality Disorder), personality constructs (such as impulsivity), and aggression. [42] Although there is substantial overlap between general externalizing and substance-related problems, there remains variance that is unique to substance problems. Environmental aspects may account for some of the unique variation in substance problems, including availability of specific substances or cultural influences on what is considered normative versus problematic use. [43]

The current data are unable to address the important question of predictive validity or utility of the multi-substance dependence construct. Future studies should examine whether issues such as clinical course and treatment outcomes may be informed by consistent versus inconsistent patterns of involvement across multiple substances. [44,45] The finding of measurement invariance across sampling schemes supports the assumption that DSM criteria are psychometrically equally appropriate for describing substance dependence in epidemiological and clinical samples. Although sample means vary, the measurement properties of DSM-IV substance dependence criteria remain fundamentally the same.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Declaration of Interest

J.D. was supported by NIH grants DA029377 and MH016880. A.A. was supported by NIH grants DA23668 and DA25886.

Minnesota Twin Family Study (MTFS). Funding support for the Minnesota Twin Family Study (MTFS) was provided by NIH grants AA09367 and DA05147

Study of Addiction: Genetics and Environment (SAGE). Funding support for the Study of Addiction: Genetics and Environment (SAGE) was provided through the NIH Genes, Environment and Health Initiative [GEI] (U01 HG004422). SAGE is one of the genome-wide association studies funded as part of the Gene Environment Association Studies (GENEVA) under GEI. Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01 HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Support for collection of datasets and samples was provided by the Collaborative Study on the Genetics of Alcoholism (COGA; U10 AA008401), the Collaborative Genetic Study of Nicotine Dependence (COGEND; P01 CA089392), and the Family Study of Cocaine Dependence (FSCD; R01 DA013423). Funding support for genotyping, which was performed at the Johns Hopkins University Center for Inherited Disease Research, was provided by the NIH GEI (U01HG004438), the National Institute on Alcohol Abuse and Alcoholism, the National Institute on Drug Abuse, and the NIH contract "High throughput genotyping for studying the genetic contributions to human disease" (HHSN268200782096C).

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, fourth edition, text revision (DSM-IV-TR). Washington, D.C.: 2000.

2. Hasin DS. Introduction to DSM-5 criteria linked papers in drug and alcohol dependence. *Drug and Alcohol Dependence*. 2012; 122:20–21. [PubMed: 22137644]
3. Helzer JE, Bucholz KK, Gossop M. A dimensional option for the diagnosis of substance dependence in DSM-V. *International Journal of Methods in Psychiatric Research*. 2007; 16:S24–S33. [PubMed: 17623392]
4. O'Brien C. Addiction and dependence in DSM-V. *Addiction*. 2011; 106:866–867. [PubMed: 21477226]
5. Dawson DA, Saha TD, Grant BF. A multidimensional assessment of the validity and utility of alcohol use disorder severity as determined by item response theory models. *Drug and Alcohol Dependence*. 2010; 107:31–38. [PubMed: 19782481]
6. Harford TC, Yi H, Faden VB, Chen CM. The dimensionality of DSM-IV alcohol use disorders among adolescent and adult drinkers and symptom patterns by age, gender, and race/ethnicity. *Alcoholism, Clinical and Experimental Research*. 2009; 33:1–11.
7. Saha TD, Chou SP, Grant BF. Toward an alcohol use disorder continuum using item response theory: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Psychological Medicine*. 2006; 36:931–941. [PubMed: 16563205]
8. Martin CS, Chung T, Kirisci L, Langenbucher JW. Item response theory analysis of diagnostic criteria for alcohol and cannabis use disorders in adolescents: implications for DSM-V. *Journal of Abnormal Psychology*. 2006; 115:807–814. [PubMed: 17100538]
9. Langenbucher JW, Labouvie E, Martin CS, Sanjuan PM, Bavly L, Kirisci L, Chung T. An application of item response theory analysis to alcohol, cannabis, and cocaine criteria in DSM-IV. *Journal of Abnormal Psychology*. 2004; 113:72–80. [PubMed: 14992659]
10. Gillespie NA, Neale MC, Prescott CA, Aggen SH, Kendler KS. Factor and item-response analysis DSM-IV criteria for abuse of and dependence on cannabis, cocaine, hallucinogens, sedatives, stimulants and opioids. *Addiction*. 2007; 102:920–930. [PubMed: 17523987]
11. Lynskey MT, Agrawal A. Psychometric properties of DSM assessments of illicit drug abuse and dependence: results from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *Psychological Medicine*. 2007; 37:1345–1355. [PubMed: 17407621]
12. Kahler CW, Strong DR, Kayaki J, Ramsey SE, Brown RA. An item response analysis of the alcohol dependence scale in treatment-seeking alcoholics. *Journal of Studies on Alcohol*. 2003; 64:127–136. [PubMed: 12608493]
13. Kahler CW, Strong DR, Read JP, Palfai TP, Wood MD. Mapping the continuum of alcohol problems in college students: a Rasch model analysis. *Psychology of Addictive Behaviors*. 2004; 18:322–333. [PubMed: 15631604]
14. Krueger RF, Nichol PE, Hicks BM, Markon KE, Patrick CJ, Iacono WG, McGue M. Using latent trait modeling to conceptualize an alcohol problems continuum. *Psychological Assessment*. 2004; 16:107–119. [PubMed: 15222807]
15. Morey L, Hopwood CJ. An IRT-based measure of alcohol trait severity and the role of traitedness in trait validity: A reanalysis of Project MATCH data. *Drug and Alcohol Dependence*. 2009; 105:177–184. [PubMed: 19665326]
16. Neal DJ, Corbin WR, Fromme K. Measurement of alcohol-related consequences among high school and college students: application of item response models to the Rutgers Alcohol Problems Index. *Psychological Assessment*. 2006; 18:402–414. [PubMed: 17154761]
17. Annaheim B, Scotto TJ, Gmel G. Revising the Cannabis Use Disorders Identification Test (CUDIT) by means of Item Response Theory. *International Journal of Methods in Psychiatric Research*. 2010; 19:142–155. [PubMed: 20812291]
18. Piontek D, Kraus L, Klempova D. Short scales to assess cannabis-related problems: a review of psychometric properties. *Substance Abuse Treatment, Prevention, and Policy*. 2008; 3:25.
19. Lennox R, Dennis ML, Scott CK, Funk R. Combining psychometric and biometric measures of substance use. *Drug and Alcohol Dependence*. 2006; 83:95–103. [PubMed: 16368199]
20. Hayatbakhsh MR, Najman JM, Jamrozik K, Al Mamun A, Bor W, Alati R. Adolescent problem behaviours predicting DSM-IV diagnoses of multiple substance use disorder: Findings of a prospective birth cohort study. *Social Psychiatry and Psychiatric Epidemiology*. 2008; 43:356–363. [PubMed: 18301851]

21. Kirisci L, Tarter RE, Vanyukov M, Martin C, Mezzich A, Brown S. Application of item response theory to quantify substance use disorder severity. *Addictive Behaviors*. 2006; 31:1035–1049. [PubMed: 16647219]
22. Derringer J, Krueger RF, Iacono WG, McGue M. Modeling the impact of age and sex on a dimension of poly-substance use in adolescence: A longitudinal study from 11- to 17-years-old. *Drug and Alcohol Dependence*. 2010; 110:193–199. [PubMed: 20456876]
23. Kirisci L, Vanyukov M, Dunn M, Tarter R. Item Response Theory models of substance use: An index based on 10 drug categories. *Psychology of Addictive Behaviors*. 2002; 16:290–298. [PubMed: 12503901]
24. Burton DB, Ryan JJ, Axelrod BN, Schellenberger T. A confirmatory factor analysis of the WAIS-III in a clinical sample with crossvalidation in the standardization sample. *Archives of Clinical Neuropsychology*. 2002; 17:371–387. [PubMed: 14589721]
25. Dickinson D, Ragland JD, Calkins ME, Gold JM, Gur RC. A comparison of cognitive structure in schizophrenia patients and healthy controls using confirmatory factor analysis. *Schizophrenia Research*. 2006; 85:20–29. [PubMed: 16616831]
26. Abdel-Khalek AM. A general factor of death distress in seven clinical and non-clinical groups. *Death Studies*. 2004; 28:889–898. [PubMed: 15493083]
27. Müller J, Bühner M, Ellgring H. Is there a reliable factorial structure in the 20-item Toronto Alexithymia Scale? A comparison of factor models in clinical and normal adult sample. *Journal of Psychosomatic Research*. 2003; 55:561–568. [PubMed: 14642988]
28. Reynolds MD, Tarter RE, Kirisci L. Characteristics of men with substance use disorder consequent to illicit drug use: comparison of a random sample and volunteers. *Drug and Alcohol Dependence*. 2004; 75:241–251. [PubMed: 15283945]
29. Berman AH, Bergman H, Palmstierna T, Schlyter F. Evaluation of the Drug Use Disorders Identification Test (DUDIT) in criminal justice and detoxification settings and in a Swedish population sample. *European Addiction Research*. 2005; 11:22–31. [PubMed: 15608468]
30. Iacono WG, McGue M. Minnesota Twin Family Study. *Twin Research and Human Genetics*. 2002; 5:482–487.
31. Iacono WG, McGue M, Krueger RF. Minnesota Center for Twin and Family Research. *Twin Research and Human Genetics*. 2006; 9:978–984. [PubMed: 17254440]
32. Bierut LJ, Madden PAF, Breslau N, Johnson EO, Hatsukami D, Pomerleau OF, Ballinger DG. Novel genes identified in a high-density genome wide association study for nicotine dependence. *Human Molecular Genetics*. 2007; 16:24–35. [PubMed: 17158188]
33. Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JI Jr, Schuckit MA. A new, semi-structured psychiatric interview for use in genetic linkage studies: A report on the reliability for the SSAGA. *Journal of Studies on Alcohol*. 1994; 55:149–158. [PubMed: 8189735]
34. Hesselbrock M, Easton C, Bucholz KK, Schuckit M, Hesselbrock V. A validity study of the SSAGA—a comparison with the SCAN. *Addiction*. 1999; 94:1361–1370. [PubMed: 10615721]
35. Cottler, LB. Composite International Diagnostic Interview—Substance Abuse Module (SAM). St. Louis, MO: Department of Psychiatry, Washington University School of Medicine; 2000.
36. Muthén, LK.; Muthén, BO. *Mplus User's Guide*. Sixth Edition. Los Angeles, CA: Muthén & Muthén; 1998–2010.
37. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6:1–55.
38. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*. 2002; 9:233–255.
39. Embretson, SE.; Reise, SP. *Item Response Theory for Psychologists*. Mahwah, New Jersey: Lawrence Erlbaum Associates; 2000.
40. Caetano R, Schafer J. DSM-IV alcohol dependence in a treatment sample of white, black, and Mexican-American men. *Alcoholism: Clinical and Experimental Research*. 1996; 20:384–390.
41. Schafer J, Caetano R. The DSM-IV construct of cocaine dependence in a treatment sample of black, Mexican American, and white men. *Psychological Assessment*. 1996; 8:304–311.

42. Krueger RF, South SC. Externalizing disorders: Cluster 5 of the proposed meta-structure for DSM-V and ICD-11. *Psychological Medicine*. 2009; 39:2061–2070. [PubMed: 19796431]
43. Degenhardt L, Chiu WT, Sampson N, Kessler RC, Anthony JC, Angermeyer M, Wells JE. Toward a global view of alcohol, tobacco, cannabis, and cocaine use: findings from the WHO World Mental Health Surveys. *PLoS Medicine*. 2008; 5(7):e141. [PubMed: 18597549]
44. Chung T, Martin CS, Clark DB. Concurrent change in alcohol and drug problems among treated adolescents over three years. *Journal of Studies on Alcohol and Drugs*. 2008; 69:420–429. [PubMed: 18432385]
45. Kahler CW, Strong DR, Papandonatos GD, Colby SM, Clark MA, Boergers J, Buka SL. Cigarette smoking and the lifetime alcohol involvement continuum. *Drug and Alcohol Dependence*. 2008; 93:111–120. [PubMed: 17964082]

Figure 1. Hierarchical model, in which ever use (Use) and dependence symptoms (denoted Sx₁–Sx₇) are indicators of dependence on each substance, which are in turn indicators of the higher-order multi-substance dependence latent trait.

Figure 2.

Criterion information curves for each substance (derived from parameters shown in Table 3*). The peak height of each criterion's curve represents the relative loading of that criterion. The horizontal location of the peak is the threshold, or the Z-score latent trait level at which the likelihood of an individual endorsing that criterion is 50%.

* Information was calculated as $I(\theta) = 1.7 \cdot \frac{P_i(\theta)^2}{P_i(\theta)(1-P_i(\theta))}$ where $P_i(\theta)$ is the normal-metric criterion discrimination (as given in Table 3) and $P_i(\theta)$ is the probability of an individual with latent trait level θ endorsing that criterion. [39]

Table 1

Ever use and dependence symptom endorsement rates by sex- and study-defined groups.

Substance	Criterion	MTFS males (N=1177)	MTFS females (N=1299)	SAGE males (N=1882)	SAGE females (N=2239)
Alcohol	Ever use	94.3%	88.8%	99.7%	99.6%
	Tolerance	52.9%	21.2%	61.1%	35.0%
	Withdrawal	17.3%	7.0%	33.5%	17.1%
	Using larger amounts or for longer than intended	39.4%	18.5%	72.5%	50.7%
	Persistent desire or unable to cut down	16.4%	8.8%	59.6%	34.7%
	Great amount of time spent obtaining/using/recovering	8.3%	3.0%	37.8%	20.0%
	Activities given up/reduced	4.7%	2.6%	39.4%	18.9%
	Continued use knowing causes physical/psychological problems	17.6%	8.5%	56.9%	36.1%
Marijuana	Ever use	60.0%	49.4%	81.8%	73.4%
	Tolerance	14.6%	6.8%	27.6%	11.2%
	Withdrawal	10.0%	4.5%	18.7%	8.5%
	Using larger amounts or for longer than intended	10.9%	4.5%	24.7%	11.9%
	Persistent desire or unable to cut down	11.7%	5.0%	26.1%	12.8%
	Great amount of time spent obtaining/using/recovering	14.5%	5.9%	29.0%	11.4%
	Activities given up/reduced	6.0%	2.3%	19.1%	6.9%
	Continued use knowing causes physical/psychological problems	13.5%	6.5%	21.9%	11.4%
Cocaine	Ever use	16.7%	8.7%	55.0%	34.7%
	Tolerance	1.9%	0.5%	29.9%	18.7%
	Withdrawal	2.4%	1.1%	31.4%	19.9%
	Using larger amounts or for longer than intended	2.3%	1.0%	34.5%	20.9%
	Persistent desire or unable to cut down	2.0%	1.0%	33.3%	20.5%
	Great amount of time spent obtaining/using/recovering	2.3%	0.5%	29.8%	18.9%
	Activities given up/reduced	0.7%	0.3%	29.2%	17.9%
	Continued use knowing causes physical/psychological problems	2.3%	1.0%	28.9%	18.5%

Table 2

Model fit statistics for the hierarchical model. Well-fitting models are those with values of CFI and TLI greater than 0.95, and RMSEA less than 0.06.

Model	N	CFI	TLI	RMSEA	# Free
Total sample	6597	0.999	0.999	0.022	51
MTFS males	1177	0.976	0.974	0.033	51
MTFS females	1299	0.961	0.957	0.038	51
SAGE males	1882	0.999	0.999	0.025	49*
SAGE females	2239	0.998	0.998	0.021	51
Multigroup	6597				
Unconstrained model		0.998	0.997	0.023	204
Constrained model		0.996	0.996	0.027	135

Free = Number of parameters freely estimated within the model.

* One criterion (cocaine 'ever use') was removed to identify the single-group model among SAGE males, due to multicollinearity issues (thus the number of parameters is reduced by one loading and one threshold).

Table 3

Criterion loadings (where discrimination is the unstandardized loading) and thresholds (the standardized (Z-score) latent trait level at which an individual is 50% likely to endorse that criterion).

Substance	Criterion	Discrimination	Threshold	Loading*
Alcohol	Ever use	0.29	−6.68	0.38
	Tolerance	0.79	0.90	0.75
	Withdrawal	0.79	2.16	0.75
	Using larger amounts or for longer than intended	0.88	0.59	0.79
	Persistent desire or unable to cut down	0.89	1.33	0.79
	Great amount of time spent obtaining/using/recovering	1.34	2.09	0.89
	Activities given up/reduced	1.72	2.12	0.93
	Continued use knowing causes physical/psychological problems	0.90	1.35	0.79
Marijuana	Ever use	0.92	−0.27	0.90
	Tolerance	0.92	2.79	0.90
	Withdrawal	0.97	3.22	0.91
	Using larger amounts or for longer than intended	0.81	2.91	0.87
	Persistent desire or unable to cut down	0.68	2.98	0.83
	Great amount of time spent obtaining/using/recovering	1.38	2.69	0.95
	Activities given up/reduced	0.85	3.40	0.88
	Continued use knowing causes physical/psychological problems	0.74	2.96	0.85
Cocaine	Ever use	1.20	2.03	0.89
	Tolerance	1.71	3.49	0.94
	Withdrawal	2.10	3.33	0.96
	Using larger amounts or for longer than intended	1.61	3.24	0.94
	Persistent desire or unable to cut down	1.88	3.29	0.95
	Great amount of time spent obtaining/using/recovering	1.80	3.44	0.95
	Activities given up/reduced	1.46	3.57	0.92
	Continued use knowing causes physical/psychological problems	1.67	3.49	0.94
Multi-substance	Alcohol	--	--	0.72
	Marijuana	--	--	0.89
	Cocaine	--	--	0.79

* Standardized to MTFS males as the reference group.