Published in final edited form as:

Psychol Addict Behav. 2012 March ; 26(1): 68–77. doi:10.1037/a0026230.

Are the symptoms of cannabis use disorder best accounted for by dimensional, categorical, or factor mixture models? A comparison of male and female young adults

Nathan A. Gillespie^{1,2}, Michael C. Neale^{1,3}, Lisa N. Legrand⁴, William G. Iacono⁴, and Matt McGue⁴

¹Virginia Institute for Psychiatric and Behavior Genetics, Department of Psychiatry, Virginia Commonwealth University

²Queensland Institute of Medical Research, Brisbane, Australia

³Department of Human Genetics, Virginia Commonwealth University

⁴Department of Psychology, University of Minnesota

Abstract

Despite the consensus that criteria for cannabis abuse and dependence and symptoms of withdrawal are best explained by a single latent liability, it remains unknown whether alternate models provide a better explanation of these criteria. A series of latent factor, latent class and hybrid factor mixture models were fitted to data from 872 recent cannabis users from the Minnesota Twin Family Study who completed DSM-III-R and DSM-IV diagnostic criteria for cannabis abuse, dependence and symptoms of withdrawal. Despite theoretical appeal, results did not support latent class or factor mixture modelling. Instead, symptoms of abuse, dependence and withdrawal were better summarized by a single latent factor ‘cannabis use disorder’ (CUD) for males and females. An almost two-fold sex difference in item endorsement was best explained by a single mean difference on the CUD factor, indicating that males have a greater latent liability towards expressing CUD.

Introduction

Cannabis is the most widely used illicit drug in developed countries including the United States (Dennis, Babor, Roebuck, & Donaldson, 2002; Hall, Johnston, & Donnelly, 1999). Population-based estimates of lifetime cannabis use in the USA between 1990 and 2004 range from 41.2% to 55.9% (Agrawal, Neale, Prescott, & Kendler, 2004b; Degenhardt et al., 2008; K. S. Kendler, Jacobson, Prescott, & Neale, 2003; K. S. Kendler, Karkowski, Neale, & Prescott, 2000; Samhsa, 1998; von Sydow et al., 2001). Estimates of the lifetime prevalence of abuse range from 5.5% to 8.4%, whereas lifetime dependence estimates range from 1.3% to 2.2% (Agrawal & Lynskey, 2007; Stinson et al., 2005; von Sydow, et al., 2001). The separation of the DSM-IV diagnostic criteria into the two distinct, categorical outcomes of abuse and dependence (American Psychiatric Association, 1980, 1987, 1994; Edwards, Arif, & Hadgson, 1981) has been challenged in recent years. A number of reports

Address for correspondence: Nathan A Gillespie, PhD, Virginia Institute for Psychiatric and Behavior Genetics, Department of Psychiatry Virginia Commonwealth University, 800 East Leigh Street, Biotech 1, Suite 101, Rm 1-119, Richmond VA 23219-1534, United States, ngillespie@vcu.edu.

Disclaimer

The content is the responsibility of the authors and does not represent the views of the US National Institute on Drug Abuse or the US National Institutes of Health.

examining cannabis and other substances using clinical and population-based samples have shown that the physiological, behavioural and cognitive components of substance abuse and dependence, including withdrawal, are more consistent with a single psychometric factor model (Feingold & Rounsaville, 1995a, 1995b; N.A. Gillespie, Neale, Prescott, Aggen, & Kendler, 2007; Hartman et al., 2008; Langenbucher et al., 2004; Lynskey & Agrawal, 2007; Nelson, Rehm, Ustun, Grant, & Chatterji, 1999; Teesson, Lynskey, Manor, & Baillie, 2002). Consequently, the DSM-V is set to merge the abuse-dependence criteria along with symptoms of withdrawal into a single cannabis use disorder (CUD).

Despite this consensus, there is an absence in the cannabis literature in which alternative explanations concerning the co-aggregation of symptoms have been tested and compared. Conceptually, LFA and LCA represent competing hypotheses regarding symptom or item covariation. Historically, they have been applied independently and for different purposes. In reality, they are not mutually exclusive because ordered discrete distributions of latent classes can be used to approximate a latent trait (Vermunt, 2004; Vermunt & Hagnaars, 2004) in the same way a histogram approximates a continuous distribution (Schmitt, 2006).

Latent factor analysis (LFA) (Spearman, 1904) predicts that correlations between diagnostic criteria can be explained in terms of a reduced number of unobserved variables called latent factors. This method has proved useful for developing, validating and calibrating drug abuse and dependence within and across substances (Harford & Muthen, 2001; Hasin et al., 1997; Kirisci et al., 2009; Lennox, Zarkin, & Bray, 1996; Nelson, et al., 1999), including cannabis (Feingold & Rounsaville, 1995a, 1995b; N.A. Gillespie, et al., 2007; Hartman, et al., 2008; Langenbucher, et al., 2004; Lynskey & Agrawal, 2007; Nelson, et al., 1999; Teesson, et al., 2002). The implicit conclusion drawn by cannabis researchers that have reported best fitting single factor structures assumes that drug-using populations are homogenous. This is premature. For instance, there are significant cohort differences in the prevalence of use when measured in the US between 1934 and 1974 (K.S. Kendler, Gardner, Jacobson, Neale, & Prescott, 2005), including higher rates of use, abuse and dependence reported among males (Agrawal & Lynskey, 2007; Agrawal, Neale, et al., 2004b; K.S. Kendler, et al., 2005; von Sydow, et al., 2001). In addition to the need to determine if these cohort and sex differences are qualitative versus quantitative in origin (Degenhardt, Lynskey, & Hall, 2000; Johnson & Gerstein, 1998; Muller & Gmel, 2002), the trajectory from cannabis initiation to regular use, abuse and dependence could conceivably vary between subjects (phenotypic heterogeneity) or the same cannabis use disorder may result from different sets of risk factors (etiological heterogeneity) (Ellickson, Martino, & Collins, 2004; Kandel & Chen, 2000; Windle & Wiesner, 2004).

An immediate alternative comparison ought to include latent class analysis (LCA) that stringently assumes no covariance between diagnostic criteria. A population is predicted to consist of a small number of classes that differ in their means or variances (or item thresholds and item response probabilities) on at least two or more criteria. This approach has been used to identify severity spectrums for alcohol (Bucholz et al., 1996; Lynskey et al., 2005) and cannabis use (Grant et al., 2006). They have also been applied to evaluate drug treatment response and evaluation (Grant, et al., 2006). Although theoretically plausible, a limitation of LCA is that very minor differences between classes can make it difficult to distinguish one class from the next. Moreover, because of the local independence assumption (Lazarsfeld & Henry, 1968), LCA, unlike LFA, LCA does not distinguish individual within-class differences in severity (Muthen, 2006).

Hybrid latent class and factor models can circumvent the limitations of LFA and LCA, which historically have been applied independently, by providing a useful bridge between the two modelling traditions in the analysis of diagnostic criteria (Dolan & Maas, 1998;

Everitt, 1988; Jedidi, Jagpal, & Desarbo, 1997; McLachlan & Peel, 2000; Muthen, 2006; Muthen & Shedden, 1999; Yung, 1997). This approach can be used to identify simultaneously categorical latent classes within a population as well as continuous variation or individual differences within classes. In factor mixture models (FMM), the assumption of no covariance between items i.e. item independence, is relaxed to allow for covariance between items but only within each class. In principle, FMMs may fit the data better than either LFA or LCA models and if so, provide a superior measurement and classification for complex phenotypes (see Helzer, van den Brink, & Guth, 2006). Simulations have demonstrated that it is possible to discriminate between latent class and factorial models using categorical data (Lubke & Neale, 2006, 2008). Recent evidence has shown that factor mixture models provide a better fit to DSM-IV criteria for prescription opioid abuse and dependence in adults than standard categorical diagnostics (Wu, Woody, Yang, Pan, & Blazer, 2010). Similarly, a factor mixture model appears to provide a good, sex-invariant fit to DSM-IV diagnostic criteria for alcohol abuse and dependence (Kuo, Aggen, Prescott, Kendler, & Neale, 2008) and heavy drinking (Muthen & Muthen, 2000).

Our aim is to provide the first omnibus comparison between competing, alternate explanation of DSM-III-R and DSM-IV cannabis diagnostic criteria in order to derive the best-fitting CUD phenotype from existing, widely used assessment items. Despite the fact that the genetic and environmental risks for cannabis use appear to be identical across sex (K.S. Kendler, et al., 2005) rates of cannabis use and abuse are higher among males (Coffey et al., 2002; Kessler et al., 1994). It is unclear if this sex difference is attributable to differences in item endorsement or reflects a greater latent liability among males indicated by latent mean or variance differences. Therefore, our second aim was to determine using tests of measurement invariance the source of any observed sex differences.

Methods

Participants

Described in detail elsewhere (Iacono, Carlson, Taylor, Elkins, & McGue, 1999), subjects for this study came from the Minnesota Twin Family Study (MTFS), which is a longitudinal study of two twin cohorts and their parents. Birth records and public databases were used to identify twins who were born in Minnesota between 1975-1984 and 1988-1994. These twins are predominantly Caucasians (98%) whose average socioeconomic status level (Hollingshead, 1975) corresponds to parental occupations such as clerical, sales industry and small-business owners. Only twins living within one day's driving distance to Minneapolis, who could be accompanied by a biological parent, and who were not mentally or physically handicapped were eligible for participation. Cohort 1 was first assessed at age 11 with follow-ups at ages 14, 17, 20, 24 and 29. Cohort 2 was first assessed at age 17 with follow-ups at ages 20, 24 and 29. These follow-ups were timed to coincide with major transitions in the lives of adolescents and young adults. Participation rates have generally exceeded 90% across all follow-up assessments. We chose to use data from the third and first follow-ups of Cohorts 1 and 2, respectively when twins were aged 20. The mean ages of onset and first symptoms of cannabis abuse are 18 and 19 years (N.A. Gillespie, Neale, & Kendler, 2009). The assessment is ideal, as twins will have passed through the peak risk period while still being young enough to accurately respond to diagnostic criteria. The mean age at assessment was 21.7 years ($\delta=0.01$). There were 570 females and 392 males, or 42.7% and 62.1% of the sample respectively, who reported lifetime cannabis use. Among them, there were 506 females (88.9%) and 366 males (93.3%) who reported using cannabis at least once within the last four years. Participants were on average over five years past the mean age of cannabis initiation ($\mu=16.2$ yrs, $\delta=2.0$), but still young enough to avoid recall bias associated with retrospective reporting at older ages (Pickles et al., 1998).

Measures & Reliability

DSM-III-R substance use disorder criteria were assessed using a modified version of the expanded Substance Abuse Module (SAM) developed by Robins et al (L. M. Robins, Babor, & Cottler, 1987) as a supplement to the World Health Organization's Composite International Diagnostic Interview (CIDI) (L. N. Robins et al., 1988).

Interviewers possessed a bachelor's or master's degrees in psychology or a related field, completed an extensive training program, and were blind to the clinical status of other family members. Coded interviews were reviewed by pairs of individuals with advanced clinical training (supervised by a Ph.D. clinical psychologist), also blind to other family diagnoses. At study intake, assessments were lifetime. Follow-up assessments covered the period since the last assessment, and when aggregated with each other along with the intake assessment, provided an opportunity to obtain lifetime assessments at the time of the age 20 follow-up (Leckman, Sholomskas, Thompson, Belanger, & Weissman, 1982). Information regarding onset and offset of specific symptoms and episodes, crucial in a developmental study such as ours, was also coded. We have examined reliability of our consensus procedure by disorder, informant, subject age, and instrument. Interviews of approximately 600 participants were reviewed by two independent, blinded pairs and produced kappas > 0.74 , with kappas for all substance use disorders (alcohol, nicotine, cannabis, amphetamine abuse or dependence) exceeding > 0.91 (Iacono, et al., 1999).

The eleven symptoms shown in Table 1 assessed four abuse (listed first) and seven dependence criteria each coded on a two-point [*Yes / No*] scale. Although our data were based on subjects who reported using cannabis at least once since the last assessment four years prior, administration of the diagnostic criteria was contingent upon five or more reported uses since last assessment. Subjects who reported using cannabis one through four times were coded as zero for each item. In order to correspond to the first DSM-IV substance abuse criterion, the item one is a manifestation of either "*used often when doing something important*" or "*stayed away from school or missed appointments because of use*". Likewise, endorsement of item six withdrawal is a manifestation of either "*felt sick when cutting down or stopped use*" or "*after not using cannabis, used to prevent sickness*" to correspond with the DSM-IV substance dependence withdrawal criteria.

Statistical Analyses

We fitted latent factor (LFA) and latent class (LCA) analyses as well as factor mixture models (FMM) directly to the raw ordinal data using the Mx software package (Neale et al, 2003). Raw ordinal data methods use both complete and incomplete data thereby avoiding listwise deletion that has the added advantage of increasing the accuracy of the item thresholds and improving the covariance estimates. In LFA, the observed items are specified as being caused by the latent factor with factor loadings f_1 - f_{11} . Residual (non-factor or specific) effects denoted as R_1 - R_{11} are also included and these generate additional variation in the observed scores. More than one common factor can be specified to allow for the possibility of more latent dimensions that can influence the observed items. Specifically, we used marginal maximum likelihood (MML) to model the cannabis diagnostic criteria and to improve computation efficiency. MML (Bock & Aitken, 1981) provides a fast means of calculating model parameters by integrating the latent factor distribution using 10-point quadrature (Neale, Aggen, Maes, Kubarych, & Schmitt, 2006). In MML the likelihood function is calculated as the product of the conditional probability for each item at a series of abscissae along the latent trait distribution. This product of one-dimensional integrals can be used because when conditioned on the factor, the items are independent. The Gaussian weighted sum of these conditional single item responses approximates the overall likelihood for an observed vector of item scores.

The second approach for characterizing individual differences and item responses was the latent class analysis (LCA) (Lazarsfeld, 1950). LCA assumes that any covariation between observed variables arises because the population consists of at least two subgroups, and these subgroups have different means for at least two of the measures. This situation is illustrated for two items in the scatter plots shown in Figure 1. Within each class, there is no correlation between items X and Y; hence, the scatter plots of each group are approximately circular. With a large mean difference, as in Figure 1a, the distinction between the groups is quite obvious, and clearly does not conform to a bivariate normal distribution. However, when the group mean differences are smaller (Figure 1b), the difference between this plot and that of a single bivariate normal distribution (Figure 1c) is much less obvious – both to the eye and to the statistical software designed to detect it.

The third approach, factor mixture modelling, assumes that a population may consist of several latent classes, any or all of which may vary on one or more underlying continua. An illustrative scatter plot of two variables that might be generated by a factor mixture model is shown in Figure 1d. Here the two classes differ not only in their means, but also in their covariances. The class with the higher means on both X and Y shows marked positive correlation, whereas the observations within the other class are essentially uncorrelated. Clearly, with multivariate data, a very broad array of possible models exists within this framework. Latent trait models are special cases with one or more factors but only a single class. Latent class models have at least two classes but no factors.

Our current analytic approach assumes that members of twin pairs are independent. Failure to take account of statistical non-independence of the twins is not expected to change the parameter estimates but may alter their confidence intervals. Non-independence of observations is rarely a problem when the group size is small; in the case of twin data, group size is at most two.

Model comparisons

Latent factor, latent class, and factor mixture models were fitted to the ordinal data and compared using omnibus fit indices. In addition to the traditional Akaike Information Criterion (AIC) (Akaike, 1987), we used the Bayesian Information Criterion (BIC) and sample-size adjusted (sBIC) (Schwarz, 1978) indices to identify the best-fitting models. Our decision to include the BIC and sBIC was based on simulations (Nylund, Asparouhov, & Muthén, 2007) that have shown how the BIC and sBIC (Schwarz, 1978) can outperform the AIC in complex structures in which symptoms have different endorsement probabilities for more than one latent class. Although the power to choose between competing latent class and factor models can vary according to the item response format, simulations based on data from the Virginia Twin Registry have shown that it is possible to choose between these models with categorical data (Lubke & Neale, 2008). With regard to the power to choose between competing categorical models, this depends on class separation and within-class sample size (Lubke & Muthén, 2007; Lubke & Neale, 2006). Lubke and Neale (Lubke & Neale, 2006) have shown that sample sizes as small as 75 subjects within a class can lead to correct model selection even if the distance between classes is so small that heterogeneity would be hard to detect. While bootstrap likelihood ratio tests may provide a better discrimination between models with different numbers of latent classes (Nylund, et al., 2007) this approach was not adopted because of the heavy computational burden. Instead, the best fitting model was the one where 2 out of 3 or all 3 indices showed the best fit.

For tests of sex differences, a hierarchical sequence of model comparisons was carried out using the definition variable option in Mx (Neale, Boker, Xie, & Maes, 2006). In addition to chi-squared tests, the AIC, BIC and sBIC were again used to determine the best-fitting sex effects models. First, depending on the best solution (LFA, LCA or FMM), a baseline model

representing complete sex measurement invariance (no differences) was fitted to the combined male and female data. Next, models with a sex effect on (i) the latent factor mean (or categorical latent factor variable), (ii) factor variance, (iii) factor loadings, and (iv) item thresholds (cut-off points) were then compared to the base. If comparisons between the baseline and sex effect models produce a significant misfit (significant misfit equals improved fit), then sex effects are deemed significant. Where there is no significant misfit, sex differences are deemed non-significant.

The lowest, most negative $-2LL$ value is indicative of the best fitting model, whereas the lowest, most negative AIC, BIC and sBIC values are indicative of model parsimony. This distinction is important because in Maximum Likelihood estimation log likelihoods can be improved by simply 'over fitting' or increasing the number of parameters. Indices of parsimony penalize models with increasing numbers of parameters thereby providing an index of each model's efficiency in terms of explaining observed data patterns. Our decision to include the BIC and sBIC was also based on simulations (Nylund, et al., 2007) that have shown how the Bayesian Information Criterion (BIC) and sample size adjusted BIC (sBIC) (Schwarz, 1978) can outperform the AIC in complex structures in which symptoms have different endorsement probabilities for more than one latent class.

Results

Item endorsements, phenotypic correlations & eigenvalues

Table 1 includes the item endorsements for each diagnostic criterion from which can be estimated the number of recent users who did not endorse a diagnostic symptom. For males and females alike, the most and least frequently endorsed items were '*using cannabis in hazardous situations*' and '*legal problems*' arising from use, respectively. Among males, the number of recent users who endorsed 0, 1, 2-3, and 4 or more diagnostic criteria were 49.7%, 18.0%, 14.8%, and 17.5% respectively. Among females, the number of recent users who endorsed 0, 1, 2-3, and 4 or more diagnostic criteria were 74.5%, 10.6%, 7.6%, and 7.4% respectively.

Polychoric correlations among the items also appear in Table 1. Except for item 8 '*Tried to cut down...stop using it*', nearly all correlations ranged from moderate to high. The first three eigenvalues for males and females, respectively, were 8.11, 0.91, 0.58 and 8.18, 3.28, 0.42, suggesting either a one- or two-factor dimensional structure to the data.

Model comparisons

As shown in Table 2, for males, a one-factor two-class FMM model provided a good fit under the AIC. However, across AIC, BIC and sBIC indices, the single factor LFA solution consistently outperformed all other models in terms of parsimony. For females, none of the FMMs provided a good fit to the data. The uncorrelated two factor solution performed well under the AIC and sBIC. However, again it was the single factor LFA solution that consistently outperformed all other models across all three parsimony indices. Factor loadings for the best-fitting single factor model appear in Table 3. Apart from the legal consequences item, the loadings on the single factor solution '*Cannabis Use Disorder*' were all high for males and females.

Modelling sex effects

Tests for measurement invariance to determine the nature and location of sex differences, if present, appear in Table 4. The single latent factor model (Model 1) was used as the baseline comparison. This was fitted to the combined male and female data without any sex effects included. Models 2, 3, 4, 5 and 6 represented sex effects on the latent factor variance, latent

factor mean, latent factor loadings and thresholds, latent factor loadings and item thresholds respectively. Models 3, 4 and 6 significantly reduced the overall 'misfit'. The significant misfit seen in Model 4 appears not to have been driven by sex effects on the factor loadings (Model 5) but rather by sex differences on the item thresholds (Model 6). Subsequent univariate analyses revealed that misfit in Model 6 was driven by one item: hazardous use. Although Models 4 and 6 each provided a marginally better fit to the data compared to Model 3, in terms of parsimony, Model 3 with fewer parameters consistently outperformed Models 4 and 6 in terms of larger negative AIC, BIC, and sBIC values. The best fitting model included a single sex effect on the latent mean. This model predicts that the latent liability to CUD is significantly lower among females.

Discussion

Omnibus comparisons between latent factor, latent class and factor mixture models revealed that a single latent factor best explained the patterns of association between the diagnostic criteria for cannabis abuse, dependence and symptoms of withdrawal. Despite their theoretical appeal, we found little consistent evidence to support population heterogeneity among cannabis users; neither latent class nor hybrid factor mixture models improved model fit. Our sample included subjects, who despite reporting cannabis use, endorsed no symptoms. Therefore, the extent to which these individuals represent a distinct mixture or class of asymptomatic individuals was not borne out in terms of a better fitting latent class or factor mixture model in which one might expect to retrieve an abstinent or very low use class. Instead, our findings are commensurate with a growing number of reports with similar results (Baillie & Teesson, 2010; Feingold & Rounsaville, 1995a, 1995b; N.A. Gillespie, et al., 2007; Hartman, et al., 2008; Langenbucher, et al., 2004; Lynskey & Agrawal, 2007; Nelson, et al., 1999; Teesson, et al., 2002) as well as with the proposed changes in the forthcoming DSM-V with regard to substance-related disorders.

This study and the comparisons between competing hypotheses illustrate a more general but important point - the same set of symptoms can be used to determine the defining properties of criterion in terms of dimensional continuous scales of liability, unobserved class membership structures or hybrid classifications, all of which would not be possible if the symptoms had been aggregated into an affected versus unaffected binary diagnostic variable. We know of one other published report comparing the fit of latent factor, latent class and factor mixture models to cannabis data. Baillie and Teesson (Baillie & Teesson, 2010) also found no improvement in fit for latent class or factor mixture models when compared to a single latent factor model for DSM-IV cannabis criteria measured on large population based Australia sample. This emerging consensus supports the proposed removal of the abuse-dependence distinction, as well as the re-introduction of cannabis withdrawal symptoms, in the forthcoming DSM-V.

Contrary to the previous uncertainty surrounding the clinical significance of withdrawal symptoms (American Psychiatric Association, 1994), there is now enough evidence from laboratory and clinical studies (Budney, Hughes, Moore, & Vandrey, 2004; Budney, Moore, Vandrey, & Hughes, 2003; A.H. Lichtman & Martin, 2002; A. H. Lichtman et al., 2005) showing that withdrawal diagnostic criteria are empirically valid. Although we did not assess symptoms unique to cannabis withdrawal and although inclusion of the withdrawal items in our analyses was not contingent upon symptoms causing significant occupational or social impairment (as proposed under the DSM-V) we nevertheless found that the withdrawal item loaded convincingly onto the CUD factor. In withdrawal-gate models the symptoms of withdrawal are central to diagnoses of dependence and as such their factor loadings are expected to load among the highest (Langenbucher et al., 2000). However, the size of the withdrawal factor loadings for males and females suggests that this item

discriminates no better or worse than other criteria. Overall, our results are thus supportive of findings identifying cannabis withdrawal symptoms while at the same suggesting that they form part of a larger, quantitative CUD factor comprising symptomatic withdrawal, abuse and dependence criteria.

Compared to females, males in our sample were almost twice as likely to endorse the withdrawal, abuse and dependence criterion. Although this pattern is well established (Blanco et al., 2008; Degenhardt, Cheng, & Anthony, 2007) this is first empirical report to test the location of sex differences among cannabis users. Although Model 3 with its single sex effect on the latent factor mean consistently outperformed other models, one item in Model 6, 'hazardous use', appears to be driving measurement non-invariance; males were more likely to have endorsed using cannabis in situations which increased their chances of getting hurt e.g. driving, using guns, knives, machinery, crossing traffic, climbing or swimming.

Several reports have speculated that a combination of societal and cultural factors, along with differences in biological responses and patterns of comorbidity with other psychiatric disorders, are responsible for the difference (Brady & Randall, 1999; Greenfield, Back, Lawson, & Brady, 2010; Tuchman, 2010). Concerning specific risk factors, one report (Agrawal, Jacobson, Prescott, & Kendler, 2004) argued that the stronger association between the personality dimension of novelty seeking and cannabis use in males is attributable to differences in the magnitude of the same genetic and environmental factors influencing personality and cannabis use across sex. Four twin studies have shown that sex differences in the genetic and environmental risks for cannabis use, abuse and dependence are also more likely to be quantitative rather than qualitative (Agrawal, Neale, Jacobson, Prescott, & Kendler, 2005; Agrawal, Neale, Prescott, & Kendler, 2004a; K. S. Kendler, et al., 2000; K. S. Kendler & Prescott, 1998). Our tests not only corroborate this conclusion but also demonstrate that the differences in item endorsements are better explained by a single mean difference on the latent factor, indicating that males have a greater overall liability towards expressing CUD.

Our best fitting model predicts that covariation between the abuse and dependence items is better explained by a single latent factor which is likely to be a combination of both genetic and environmental risks. This conclusion is supported by genetically informative designs which have explored the genetic and environmental etiology of cannabis use, abuse and dependence (N. A. Gillespie, Kendler, & Neale, 2011; N.A. Gillespie, et al., 2009; K. S. Kendler, et al., 2000; K. S. Kendler & Prescott, 1998). Nevertheless, our model fitting was by no means exhaustive. Item response theory (IRT) (Lord, 1952, 1980), multiple indicators multiple causes (Finch, 2005), Mantel-Hansel (Holland & Thayer, 1988), logistic regression (Swaminathan & Roger, 1990), and the likelihood ratio test (Thissen, Steinberg, & Wainer, 1993) among other methods represent various means for detecting differential item functioning (DIF). As discussed by Takane and DeLeeuw (Takane & Leeuw, 1987), there is a direct equivalence of the normal ogive item response model and factor analysis model for binary data that we have fitted. Therefore, in terms of IRT, item difficulty parameters correspond to our item thresholds (which can be derived from item endorsements in Table 1) while the discrimination parameters correspond to the observed factor loadings in Table 3. Examining these properties and detecting differential item functioning can improve scale construction through the elimination of items that are redundant or that fail to discriminate adequately. For example, the lower factor loadings, especially among females, for legal consequences "Did your use of cannabis lead to trouble with police?" suggests that this infrequently endorsed item is a poor discriminator. Indeed, based on results from a number of reports, some including formal tests of DIF (N.A. Gillespie, et al., 2007; Hartman, et al., 2008; Hasin, Paykin, Meydan, & Grant, 2000; Langenbucher, et al., 2004; Lynskey &

Agrawal, 2007; Schuckit et al., 1999; Teesson, et al., 2002) the same conclusion can be made thereby empirically justifying the removal of legal consequences from the proposed DSM-V.

Limitations

Our findings must be interpreted in the context of three important limitations. First, while the sample is representative of the state of Minnesota when subjects were born our data were restricted to Caucasian male and female twins and cannot be extrapolated to ethnic minorities. Nevertheless, our estimates of use including the prevalence of item endorsements are very similar to those found in other large, population based twin samples (N.A. Gillespie, et al., 2007) which in turn are typical of the general population of North American Caucasians in rates of psychopathologic conditions, including illicit substances (K. S. Kendler, et al., 2000). Second, the mean age at assessment was five years past the mean age of cannabis initiation. Although this is an ideal assessment age for studying cannabis use, abuse and dependence because most twins will have passed through the peak risk period while still being young enough to accurately respond to diagnostic criteria, the results cannot preclude the possibility of different latent factor or class structures at older ages performing better. However, the fact that a uni-dimensional factor structure is more often found in large population based studies with varying sample ages lends support to these findings. Third, the models fitted were not exhaustive. Our results imply that the linear coefficients (i.e. intercepts and factor loadings) that link the observed items to the latent factor are fixed or common for all subjects which may be too restrictive if subjects habitually respond to questionnaires idiosyncratically (Maydeu-Olivares & Coffman, 2006). Therefore, to test for idiosyncratic reporting we incorporated into the intercept of the model an additional latent 'random intercept factor' (RIF) described by Maydeu-Olivares and Coffman (Maydeu-Olivares & Coffman, 2006). Normally, factor loadings from the RIF are constrained to one and the RIF variance estimated. However, since we are modelling ordinal data and latent factor variance is by default one, we constrained the factor loadings for the random intercept factor to be equal. For males, the addition of the RIF did not yield a better fit (AIC= -5759.21, BIC= -10634.14, & sBIC= -4330.17) whereas for females, the RIF model provided a marginally better fit across all three indices (AIC= -9287.31, BIC= -16236.37, & sBIC= -7514.00). Although this suggest that females may have an idiosyncratic response pattern, all factor loadings remained significant with an average change of only 0.02 across items.

Conclusion

Results from an omnibus comparison between theoretically plausible and alternative measurement models revealed that the diagnostic criteria for cannabis abuse, dependence and symptoms of withdrawal were best explained by a single latent factor we labelled 'cannabis use disorder.' This was true for both males and females. The almost two-fold sex difference in item endorsement was best explained by a single mean difference on the latent factor. Males appear to have a greater latent liability towards expressing CUD. Our next goal will be to explore the genetic and environmental etiology of this CUD phenotype using biometrical genetic analyses. Derivation of the best-fitting genetic and environmental model based on CUD criteria will enable estimation of individual latent trait scores. Therefore, given genetically informative data such as ours, it will be possible to identify maximally heritable factor scores, which in turn will increase the power to detect quantitative trait loci signals for CUD as part of anticipated genome wide association analyses.

Acknowledgments

Funding was received from the US National Institute on Drug Abuse (R00DA023549, 1K99DA023549-01A2, DA-18673) and from Minnesota Center for Twin and Family Research grants (DA05147 and AA09367).

References

- Agrawal A, Jacobson KC, Prescott CA, Kendler KS. A twin study of personality and illicit drug use and abuse/dependence. *Twin Research*. 2004; 7(1):72. [PubMed: 15053856]
- Agrawal A, Lynskey MT. Does gender contribute to heterogeneity in criteria for cannabis abuse and dependence? Results from the National Epidemiological Survey on Alcohol and Related Conditions. *Drug and Alcohol Dependence*. 2007; 88:300–307. [PubMed: 17084563]
- Agrawal A, Neale MC, Jacobson KC, Prescott CA, Kendler KS. Illicit drug use and abuse/dependence: modeling of two-stage variables using the CCC approach. *Addictive Behaviors*. 2005; 30(5):1043–1048. [PubMed: 15893102]
- Agrawal A, Neale MC, Prescott CA, Kendler KS. Cannabis and other illicit drugs: comorbid use and abuse/dependence in males and females. *Behavior Genetics*. 2004a; 34(3):217–228. [PubMed: 14990863]
- Agrawal A, Neale MC, Prescott CA, Kendler KS. A twin study of early cannabis use and subsequent use and abuse/dependence of other illicit drugs. *Psychological Medicine*. 2004b; 34(7):1227–1237. [PubMed: 15697049]
- Akaike H. Factor analysis and AIC. *Psychometrika*. 1987; 52:317–332.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 3rd ed.. American Psychiatric Association; Washington, DC: 1980.
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders: DSM-III-R*. 3rd Revised ed.. American Psychiatric Association; Washington, DC: 1987.
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders: DSM-IV*. 4th ed.. American Psychiatric Association; Washington, DC: 1994.
- Baillie AJ, Teesson M. Continuous, categorical and mixture models of DSM-IV alcohol and cannabis use disorders in the Australian community. *Addiction*. 2010; 105(7):1246–1253. [PubMed: 20491729]
- Blanco C, Ogburn E, Pérez de Los Cobos J, Lujan J, Nunes EV, Grant B, Hasin DS. DSM-IV criteria-based clinical subtypes of cannabis use disorders: results from the National Epidemiological Survey on Alcohol and Related Conditions (NESARC). *Drug and Alcohol Dependence*. 2008; 96(1-2):136–144. [PubMed: 18420357]
- Bock RD, Aitken M. Marginal maximum likelihood estimation of item parameters: An application of the EM algorithm. *Psychometrika*. 1981; 46:443–460.
- Brady K, Randall CL. Gender differences in substance use disorders. *The Psychiatric Clinics of North America*. 1999; 22(2):241–252. [PubMed: 10385931]
- Bucholz KK, Heath AC, Reich T, Hesselbrock VM, Kramer JR, Nurnberger JI Jr. Schuckit MA. Can we subtype alcoholism? A latent class analysis of data from relatives of alcoholics in a multicenter family study of alcoholism. *Alcoholism: Clinical and Experimental Research*. 1996; 20(8):1462–1471.
- Budney AJ, Hughes JR, Moore BA, Vandrey R. Review of the validity and significance of cannabis withdrawal syndrome. *American Journal of Psychiatry*. 2004; 161(11):1967–1977. [PubMed: 15514394]
- Budney AJ, Moore BA, Vandrey RG, Hughes JR. The time course and significance of cannabis withdrawal. *Journal of Abnormal Psychology*. 2003; 112(3):393–402. [PubMed: 12943018]
- Coffey C, Carlin JB, Degenhardt L, Lynskey M, Sanci L, Patton GC. Cannabis dependence in young adults: an Australian population study. *Addiction*. 2002; 97(2):187–194. [PubMed: 11860390]
- Degenhardt L, Cheng H, Anthony JC. Assessing cannabis dependence in community surveys: methodological issues. *International Journal of Methods in Psychiatric Research*. 2007; 16(2):43–51. [PubMed: 17623384]

- 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]
- Degenhardt L, Lynskey M, Hall W. Cohort trends in the age of initiation of drug use in Australia. *Australian and New Zealand Journal of Public Health*. 2000; 24(4):421–426. [PubMed: 11011471]
- Dennis M, Babor TF, Roebuck MC, Donaldson J. Changing the focus: the case for recognizing and treating cannabis use disorders. *Addiction*. 2002; 97(Supplement 1):4–15. [PubMed: 12460125]
- Dolan CV, Maas H. L. J. v. d. Fitting multivariate normal finite mixtures subject to structural equation modeling. *Psychometrika*. 1998; 63:227–253.
- Edwards G, Arif A, Hadgson R. Nomenclature and classification of drug- and alcohol-related problems: a WHO Memorandum. *Bulletin of the World Health Organization*. 1981; 59(2):225–242. [PubMed: 6972816]
- Ellickson PL, Martino SC, Collins RL. Marijuana use from adolescence to young adulthood: multiple developmental trajectories and their associated outcomes. *Health Psychology*. 2004; 23(3):299–307. [PubMed: 15099171]
- Everitt BS. A Finite Mixture Model For The Clustering Of Mixed-Mode Data. *Statistics & Probability Letters*. 1988; 6:305–309.
- Feingold A, Rounsaville B. Construct validity of the abuse-dependence distinction as measured by DSM-IV criteria for different psychoactive substances. *Drug and Alcohol Dependence*. 1995a; 39(2):99–109. [PubMed: 8529538]
- Feingold A, Rounsaville B. Construct validity of the dependence syndrome as measure by DSM-IV for different psychoactive substances. *Addiction*. 1995b; 90(12):1661–1669. [PubMed: 8555957]
- Finch H. The MIMIC model as a method for detecting DIF: Comparison with Mantel-Haenszel, SIBTEST and the IRT likelihood ratio test. *Applied Psychological Measurement*. 2005; 29:278–295.
- Gillespie NA, Kendler KS, Neale MC. Psychometric modeling of cannabis initiation and use and the symptoms of cannabis abuse, dependence and withdrawal in a sample of male and female twins. *Drug and alcohol dependence*. 2011 Epub ahead of print.
- Gillespie NA, Neale MC, Kendler KS. Pathways to cannabis abuse: A multi-stage model from cannabis availability, cannabis initiation, and progression to abuse. *Addiction*. 2009; 104(3):430–438. [PubMed: 19207351]
- Gillespie NA, Neale MC, Prescott CA, Aggen SH, Kendler KS. Factor and Item-Response Analysis of DSM-IV Criteria for Abuse of and Dependence on Cannabis, Cocaine, Hallucinogens, Sedatives, Stimulants and Opioids. *Addiction*. 2007; 102:920–930. [PubMed: 17523987]
- Grant JD, Scherrer JF, Neuman RJ, Todorov AA, Price RK, Bucholz KK. A comparison of the latent class structure of cannabis problems among adult men and women who have used cannabis repeatedly. *Addiction*. 2006; 101(8):1133–1142. [PubMed: 16869843]
- Greenfield SF, Back SE, Lawson K, Brady KT. Substance abuse in women. *Psychiatr Clin North Am*. 2010; 33(2):339–355. [PubMed: 20385341]
- Hall, W.; Johnston, L.; Donnelly, N. The epidemiology of cannabis use and its consequences. In: Kalant, H.; Corrigal, W.; Hall, W.; Smart, R., editors. *The Health Effects of Cannabis*. Addiction Research Foundation; Toronto: 1999. p. 69-125.
- Harford TC, Muthen BO. The dimensionality of alcohol abuse and dependence: a multivariate analysis of DSM-IV symptom items in the National Longitudinal Survey of Youth. *Journal of Studies on Alcohol and Drugs*. 2001; 62(2):150–157.
- Hartman CA, Gelhorn H, Crowley TJ, Sakai JT, Stallings M, Young SE, Hopfer CJ. Item response theory analysis of DSM-IV cannabis abuse and dependence criteria in adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2008; 47(2):165–173. [PubMed: 18176333]
- Hasin D, Grant BF, Cottier L, Blaine J, Towle L, Ustun B, Sartorius N. Nosological comparisons of alcohol and drug diagnoses: a multisite, multi-instrument international study. *Drug and Alcohol Dependence*. 1997; 25:217–226. [PubMed: 9306047]
- Hasin D, Paykin A, Meydan J, Grant B. Withdrawal and tolerance: prognostic significance in DSM-IV alcohol dependence. *Journal of Studies on Alcohol*. 2000; 61(3):431–438. [PubMed: 10807215]

- Helzer JE, van den Brink W, Guth SE. Should there be both categorical and dimensional criteria for the substance use disorders in DSM-V? *Addiction*. 2006; 101(1):17–22. [PubMed: 16930157]
- Holland, WP.; Thayer, DT. Differential item performance and the Mantel-Haenszel procedure. In: Wainer, H.; Braun, HE., editors. *Test validity*. Lawrence Erlbaum; Hillsdale, NJ: 1988. p. 129-145.
- Hollingshead, AB. Four factor index of social status. Unpublished manuscript. Yale University; New Haven, CT: 1975.
- Iacono WG, Carlson SR, Taylor J, Elkins IJ, McGue M. Behavioral disinhibition and the development of substance use disorders: Findings from the Minnesota Twin Family Study. *Development and Psychopathology*. 1999; 11(4):869–900. [PubMed: 10624730]
- Jedidi K, Jagpal HS, Desarbo WS. Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science*. 1997; 16:39–59.
- Johnson RA, Gerstein DR. Initiation of use of alcohol, cigarettes, marijuana, cocaine, and other substances in US birth cohorts since 1919. *American Journal of Public Health*. 1998; 88(1):27–33. [PubMed: 9584029]
- Kandel DB, Chen K. Types of marijuana users by longitudinal course. *Journal of Studies on Alcohol*. 2000; 61(3):367–378. [PubMed: 10807207]
- Kendler KS, Gardner C, Jacobson KC, Neale MC, Prescott CA. Genetic and environmental influences on illicit drug use and tobacco use across birth cohorts. *Psychological Medicine*. 2005; 35(9): 1349–1356. [PubMed: 16168157]
- Kendler KS, Jacobson KC, Prescott CA, Neale MC. Specificity of genetic and environmental risk factors for use and abuse/dependence of cannabis, cocaine, hallucinogens, sedatives, stimulants, and opiates in male twins. *American Journal of Psychiatry*. 2003; 160(4):687–495. [PubMed: 12668357]
- 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. *Archives of General Psychiatry*. 2000; 57(3):261–269. [PubMed: 10711912]
- Kendler KS, Prescott CA. Cannabis use, abuse, and dependence in a population-based sample of female twins. *American Journal of Psychiatry*. 1998; 155(8):1016–1022. [PubMed: 9699687]
- Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, Kendler KS. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry*. 1994; 51(1):8–19. [PubMed: 8279933]
- Kirisci L, Tarter R, Mezzich A, Ridenour T, Reynolds M, Vanyukov M. Prediction of cannabis use disorder between boyhood and young adulthood: clarifying the phenotype and environment. *American Journal of Addiction*. 2009; 18(1):36–47.
- Kuo PH, Aggen SH, Prescott CA, Kendler KS, Neale MC. Using a factor mixture modeling approach in alcohol dependence in a general population sample. *Drug and Alcohol Dependence*. 2008; 98:105–114. [PubMed: 18586414]
- 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(1):72–80. [PubMed: 14992659]
- Langenbucher JW, Martin CS, Labouvie E, Sanjuan PM, Bavly L, Pollock NK. Toward the DSM-V: the Withdrawal-Gate Model versus the DSM-IV in the diagnosis of alcohol abuse and dependence. *Journal of Consulting and Clinical Psychology*. 2000; 68(5):799–809. [PubMed: 11068966]
- Lazarsfeld, PF. The logical and mathematical foundations of latent structure analysis. In: Stouffer, SA., editor. *The American soldier: Studies in social psychology in World War II*. Vol. 4. Princeton University Press; 1950.
- Lazarsfeld, PF.; Henry, NW. Houghton Mifflin; Boston: 1968.
- Leckman JF, Sholomskas D, Thompson WD, Belanger A, Weissman MM. Best estimate of lifetime psychiatric diagnosis: a methodological study. *Archives of General Psychiatry*. 1982; 39(8):879. [PubMed: 7103676]
- Lennox RD, Zarkin GA, Bray JW. Latent variable models of alcohol-related constructs. *Journal of Substance Abuse*. 1996; 8(2):241–250. [PubMed: 8880663]

- Lichtman AH, Martin BR. Marijuana withdrawal syndrome in the animal model. *Journal of Clinical Pharmacology*. 2002; 42(11 Suppl):20S–27S. [PubMed: 12412832]
- Lichtman AH, Martin BR, Budney AJ, Hughes JR, Moore BA, Vandrey R. Cannabinoid tolerance and dependence: Review of the validity and significance of cannabis withdrawal syndrome. *Handbook of Experimental Pharmacology*. 2005; 161(168):691–717. [PubMed: 16596793]
- Lord FM. A theory of test scores. *Psychometric Monograph*. 1952; 7
- Lord, FM. Applications of item response theory to practical testing problems. L. Erlbaum Associates; Hillsdale, N.J: 1980.
- Lubke G, Neale MC. Distinguishing Between Latent Classes and Continuous Factors: Resolution by Maximum Likelihood? *Multivariate Behavioral Research*. 2006; 41(4):499–532.
- Lubke G, Neale MC. Distinguishing between latent classes and continuous factors with categorical outcomes: Class invariance of parameters of factor mixture models. *Multivariate Behavioral Research*. 2008; 43(4):592–620. [PubMed: 20165736]
- 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(9):1345–1355. [PubMed: 17407621]
- Lynskey MT, Nelson EC, Neuman RJ, Bucholz KK, Madden PA, Knopik VS, Heath AC. Limitations of DSM-IV operationalizations of alcohol abuse and dependence in a sample of Australian twins. *Twin Research and Human Genetics*. 2005; 8(6):574–584. [PubMed: 16354499]
- Maydeu-Olivares A, Coffman DL. Random intercept item factor analysis. *Psychological methods*. 2006; 11(4):344–362. [PubMed: 17154751]
- McLachlan, GJ.; Peel, D. Finite mixture models. Wiley; New York: 2000.
- Muller S, Gmel G. Changes in the age of onset of cannabis use: results of the 2nd Swiss Health Survey 1997. *Soz Präventivmed*. 2002; 47(1):14–23. [PubMed: 12050925]
- Muthen B. Should substance use disorders be considered as categorical or dimensional? *Addiction*. 2006; 101(Suppl 1):6–16. [PubMed: 16930156]
- Muthen B, Muthen LK. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcoholism, Clinical and Experimental Research*. 2000; 24(6):882–891.
- Muthen B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*. 1999; 55:463–469. [PubMed: 11318201]
- Neale MC, Aggen SH, Maes HH, Kubarych TS, Schmitt JE. Methodological issues in the assessment of substance use phenotypes. *Addictive Behaviors*. 2006; 31(6):1010–1034. [PubMed: 16723188]
- Neale, MC.; Boker, SM.; Xie, G.; Maes, HH. Mx: Statistical Modeling. 7th Edition. Department of Psychiatry; Box 126 MCV, Richmond, VA 23298: 2006.
- Nelson CB, Rehm J, Ustun TB, Grant B, Chatterji S. Factor structures for DSM-IV substance disorder criteria endorsed by alcohol, cannabis, cocaine and opiate users: results from the WHO reliability and validity study. *Addiction*. 1999; 94(6):843–855. [PubMed: 10665074]
- Nylund KL, Asparouhov T, Muthén B. Deciding on the number of classes in latent class analysis and growth mixture modeling. A Monte Carlo simulation study. *Structural Equation Modeling*. 2007; 14:535–569.
- Pickles A, Pickering K, Simonoff E, Silberg J, Meyer J, Maes H. Genetic ‘Clocks’ and ‘Soft’ Events: A Twin Model for Pubertal Development and Other Recalled Sequences of Developmental Milestones, Transitions, or Ages at Onset. *Behavior Genetics*. 1998; 28(4):243–253. [PubMed: 9803017]
- Robins, LM.; Babor, TF.; Cottler, LB. Composite international diagnostic interview: expanded substance abuse module. Washington University Department of Psychiatry; St. Louis, MO: 1987.
- Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke J. The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry*. 1988; 45(12):1069–1077. [PubMed: 2848472]
- Samhsa. National Household Survey on Drug Abuse Main Findings 1996. National Clearinghouse for Alcohol & Drug Info (NCADI); Rockville, MD: 1998.

- Schuckit MA, Daeppen JB, Danko GP, Tripp ML, Smith TL, Li TK, Bucholz KK. Clinical implications for four drugs of the DSM-IV distinction between substance dependence with and without a physiological component. *American Journal of Psychiatry*. 1999; 156(1):41–49. [PubMed: 9892296]
- Schwarz G. Estimating the dimension of a model. *Annals of Statistics*. 1978; 6:461–464.
- Stinson FS, Grant BF, Dawson DA, Ruan WJ, Huang B, Saha T. Comorbidity between DSM-IV alcohol and specific drug use disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Drug and Alcohol Dependence*. 2005; 80(1):105–116. [PubMed: 16157233]
- Swaminathan H, Roger HJ. Detecting differential item functioning using logistic regression procedure. *Journal of Educational Measurement*. 1990; 27(4):361–370.
- Takane Y, Leeuw J. d. On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika*. 1987; 52(3):393–408.
- Teesson M, Lynskey M, Manor B, Baillie A. The structure of cannabis dependence in the community. *Drug and Alcohol Dependence*. 2002; 68(3):255–262. [PubMed: 12393220]
- Thissen, D.; Steinberg, L.; Wainer, H. Detection of differential item functioning using the parameters of item response models. In: Holland, PW.; Wainer, H., editors. *Differential Item Functioning*. Lawrence Erlbaum Associates; Hillsdale, NJ: 1993. p. 67-113.
- Tuchman E. Women and addiction: the importance of gender issues in substance abuse research. *Journal of Addictive Diseases*. 2010; 29(2):127–138. [PubMed: 20407972]
- von Sydow K, Lieb R, Pfister H, Hofler M, Sonntag H, Wittchen HU. The natural course of cannabis use, abuse and dependence over four years: a longitudinal community study of adolescents and young adults. *Drug and Alcohol Dependence*. 2001; 64(3):347–361. [PubMed: 11672949]
- Windle M, Wiesner M. Trajectories of marijuana use from adolescence to young adulthood: predictors and outcomes. *Developmental Psychopathology*. 2004; 16(4):1007–1027.
- Wu LT, Woody GE, Yang C, Pan JJ, Blazer DG. Abuse and dependence on prescription opioids in adults: a mixture categorical and dimensional approach to diagnostic classification. *Psychological Medicine*. 2010; 41:653–664. [PubMed: 20459887]
- Yung YF. Finite mixtures in confirmatory factor-analysis models. *Psychometrika*. 1997; 62:297–330.

Figure 1.

Hypothetical scatter plots of two items to illustrate expected differences in distributions between latent factor, latent class and factor mixture models: (a) latent class model with large mean differences; (b) latent class model with small mean differences; (c) latent factor model with correlated items; and (d) factor mixture model (FMM). In the FMM, the group with higher means (red, upper right) shows positive within-group correlation whereas the group with lower means (green, lower left) does not.

Table 1

Item endorsements¹ (%) for the abuse and dependence diagnostic criteria based on subjects who reported having tried cannabis. Also included are phenotypic polychoric² correlations for male (below diagonal) and female cannabis users.

Diagnostic criteria	Endorsements		Item correlation										
	Males	Females	1	2	3	4	5	6	7	8	9	10	11
1. Used often when doing something...important / fail to fulfil obligations ³	14.3%	6.8%	1	.76	.32	.79	.80	.84	.72	.84	.79	.84	.72
2. Hazardous use	44.9%	19.0%	.77	1	.63	.72	.74	.70	.47	.65	.77	.63	.53
3. Consequences: legal	1.6%	0.2%	.64	.47	1	.69	.40	.25	.13	.48	.52	.39	.17
4. Consequences: social	9.6%	3.8%	.62	.79	.74	1	.70	.62	.58	.75	.78	.80	.57
5. Need for larger amounts/doses (tolerance)	18.5%	10.6%	.67	.73	.29	.65	1	.79	.63	.80	.81	.92	.70
6. Withdrawal symptoms: feeling sick when cutting down / stopping ⁴	10.7%	4.6%	.67	.60	.51	.54	.71	1	.63	.84	.78	.72	.62
7. Used more or longer than thought/planned	9.3%	4.8%	.66	.74	.25	.50	.74	.64	1	.64	.66	.78	.56
8. Loss of control: unable to stop / desire to stop, tried to cut down / stop using it 11.6%		4.8%	.61	.61	.45	.56	.71	.65	.74	1	.72	.83	.71
9. Spend time taking / using it, recovering from it, or doing whatever	16.2%	8.6%	.73	.75	.56	.74	.75	.69	.72	.65	1	.83	.55
10. Used instead of work/hobbies	5.8%	2.4%	.83	.87	.58	.69	.75	.73	.81	.67	.74	1	.77
11. Consequences: physical & psychological	11.0%	5.0%	.75	.53	.51	.48	.65	.66	.62	.58	.69	.75	1

Note: Endorsements and item correlations based on 506 females and 366 males who reported using cannabis at least once within the last four years.

¹ Item endorsements represent the standard normal cumulative distribution based on marginal maximum likelihood threshold estimates. The estimated proportion of individuals who endorse an item when calculated under MML will differ from the sample statistic because the location of item's threshold is sensitive to the correlation between diagnostic criteria.

² Based on the Central Limit Theorem polychoric correlations assume that underlying each ordinal response there is continuously and normally distributed scale of liability which is multi-factorial and that the joint distribution of this scale with liability scales underlying other ordinal items is bivariate normal.

³ Manifestation of either "used often when doing something important" or "stayed away from school or missed appointments because of use" to correspond with DSM-IV substance abuse, criteria A1.

⁴ Withdrawal as manifested by either "felt sick when cutting down or stopped use" or "after not using cannabis, used to prevent sickness" to correspond with DSM-IV substance dependence withdrawal criteria.

Table 2

Comparison of the empirical, latent factor (LFA), latent class (LCA) and factor mixture models (FMM) by way of Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and sample size adjusted Bayesian Information Criterion (sBIC). LFA modelling also included one 2-factor oblique solution. Under each fit index the two best-fitting models are bolded.

Males	-2LL	df	AIC	BIC	sBIC
LFA - 1 dimension	2190.71	3975	-5759.29	-10636.13	-4330.58
LFA - 2 dimensions	2173.59	3964	-5754.41	-10612.23	-4324.12
LFA - 2 correlated dimensions	2169.76	3963	-5756.24	-10611.19	-4324.68
LFA - 3 dimensions	2163.34	3953	-5742.66	-10584.89	-4314.23
LCA - 1 class	2904.99	3987	-5069.01	-10314.41	-3989.82
LCA - 2 classes	2240.87	3974	-5707.14	-10608.10	-4304.14
LCA - 3 classes	2177.00	3962	-5747.00	-10604.62	-4319.69
FMM - 1 factor 2 classes	2158.71	3952	-5745.29	-10584.25	-4315.18
FMM - 1 factor 3 classes	2112.80	3929	-5745.20	-10539.32	-4306.74
Females					
LFA - 1 dimension	1712.69	5497	-9281.31	-16235.48	-7511.52
LFA - 2 dimensions	1692.32	5486	-9279.68	-16211.46	-7504.97
LFA - 2 correlated dimensions	1713.97	5485	-9256.03	-16197.53	-7492.62
LFA - 3 dimensions	1894.71	5475	-9055.29	-16076.06	-7387.03
LCA - 1 class	2427.05	5509	-8590.95	-15915.61	-7172.61
LCA - 2 classes	1750.88	5496	-9241.12	-16213.27	-7490.91
LCA - 3 classes	1690.96	5484	-9277.04	-16205.92	-7502.60
FMM - 1 factor 2 classes	1672.12	5474	-9275.88	-16184.25	-7496.80
FMM - 1 factor 3 classes	1658.92	5451	-9243.08	-16119.34	-7468.39

-2LL = -2xLog Likelihood

Table 3

Factor loadings for the best-fitting 'Cannabis Use Disorder' latent factor for males and females.

Diagnostic criteria	M	F
1. Used often when doing something...important / fail to fulfil obligations	.82	.90
2. Hazardous use	.83	.83
3. Consequences: legal	.65	.57
4. Consequences: social	.76	.84
5. Need for larger amounts/doses (tolerance)	.85	.92
6. Withdrawal symptoms: feeling sick when cutting down / stopping	.78	.87
7. Used more or longer than thought/planned	.82	.75
8. Loss of control: unable to stop / desire to stop, tried to cut down / stop using it	.75	.89
9. Spend time taking / using it, recovering from it, or doing whatever	.87	.88
10. Used instead of work/hobbies	.99	.99
11. Consequences: physical & psychological	.77	.77

Tests of measurement invariance across sex for the Minnesota cannabis abuse and dependence items. The baseline model (1) which includes no sex effects is compared to models with covariate sex effects on the (2) latent factor mean, (3) latent factor variance, (4) latent factor loadings, and (5) all item thresholds.

Table 4

Model	-2LL	df	Δ -2LL	Δ df	p	Parameters	AIC	BIC	sBIC
1. Complete invariance	3988.42	9495					-15001.58	-30128.29	-15051.40
2. Sex on LF variance	3985.00	9494	3.42	1	0.06	23	-15003.00	-30126.61	-15051.32
3. Sex on LF mean	3940.06	9494	48.36	1	***	23	-15047.94	-30149.08	-15073.79
4. Sex on LF loadings & thresholds	3903.40	9473	85.02	22	***	44	-15042.60	-30096.37	15054.42
5. Sex on LF loadings	3974.60	9484	13.82	11	0.24	33	-14993.41	-30097.98	-15038.57
6. Sex on item thresholds	3917.29	9484	71.13	11	***	33	-15050.71	-30126.64	-15067.22

-2LL = $-2 \times \log$ Likelihood, Δ -2LL = change in -2LL which is asymptotically distributed as a chi-square with the degrees of freedom equalling the difference between the fully saturated model (Model 1) with the each of nested sub-models (Model 2-5), AIC = Akaike Information Criterion, BIC = Bayesian Information Criterion, sBIC = sample-size adjusted BIC