Published in final edited form as:
Twin Res Hum Genet. 2014 April ; 17(2): 89–98. doi:10.1017/thg.2014.9.

Comparing Factor, Class, and Mixture Models of Cannabis Initiation and DSM Cannabis Use Disorder Criteria, Including Craving, in the Brisbane Longitudinal Twin Study

Thomas S. Kubarych[1], **Kenneth S. Kendler**[1,2], **Steven H. Aggen**[1], **Ryne Estabrook**[1], **Alexis C. Edwards**[1], **Shaunna L. Clark**[3], **Nicholas G. Martin**[4], **Ian B. Hickie**[5], **Michael C. Neale**[1,2], and **Nathan A. Gillespie**[1,4]

[1]Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA

[2]Department of Human Genetics, Virginia Commonwealth University, 1101 East Marshall Street, Richmond, VA, USA

[3]School of Pharmacy, Center for Biomarker Research and Personalized Medicine, Virginia Commonwealth University, 1112 East Clay Street, McGuire Hall, Richmond, VA, USA

[4]Queensland Institute of Medical Research, 300 Herston Rd, Brisbane, QLD, Australia

[5]Brain & Mind Research Institute, University of Sydney, 100 Mallett St, Camperdown, NSW, Australia

Abstract

Accumulating evidence suggests that the Diagnostic and Statistical Manual of Mental Disorders (DSM) diagnostic criteria for cannabis abuse and dependence are best represented by a single underlying factor. However, it remains possible that models with additional factors, or latent class models or hybrid models, may better explain the data. Using structured interviews, 626 adult male and female twins provided complete data on symptoms of cannabis abuse and dependence, plus a craving criterion. We compared latent factor analysis, latent class analysis, and factor mixture modeling using normal theory marginal maximum likelihood for ordinal data. Our aim was to derive a parsimonious, best-fitting cannabis use disorder (CUD) phenotype based on DSM-IV criteria and determine whether DSM-5 craving loads onto a general factor. When compared with latent class and mixture models, factor models provided a better fit to the data. When conditioned on initiation and cannabis use, the association between criteria for abuse, dependence, withdrawal, and craving were best explained by two correlated latent factors for males and females: a general risk factor to CUD and a factor capturing the symptoms of social and occupational impairment as a consequence of frequent use. Secondary analyses revealed a modest increase in the prevalence of DSM-5 CUD compared with DSM-IV cannabis abuse or dependence. It is concluded that, in addition to a general factor with loadings on cannabis use and symptoms of abuse, dependence,

ADDRESS FOR CORRESPONDENCE: Thomas S. Kubarych, Department of Psychiatry, Virginia Institute for Psychiatric and Behavior Genetics, Virginia Commonwealth University, 800 East Leigh Street, Biotech 1, Suite 1-127, Richmond VA 23219-1534, USA. tkubarych@gmail.com.

withdrawal, and craving, a second clinically relevant factor defined by features of social and occupational impairment was also found for frequent cannabis use.

Keywords

cannabis use disorder; DSM-IV; DSM-5; craving; latent class analysis; mixture modeling; factor analysis; marijuana; symptoms; abuse; dependence

---

Cannabis is the most widely used illicit drug in developed countries (Dennis et al., 2002;Hall et al., 1999). Population-based estimates of lifetime cannabis use in the United States between 1990 and 2004 range from 41.2% to 55.9% (Agrawal & Lynskey, 2007; American Psychiatric Association, 1980, 1987, 1994; Edwards et al., 1981; Stinson et al., 2005; von Sydow et al., 2001). For cannabis use disorder (CUD) in the United States, rates of lifetime abuse range from 5.5% to 8.4%, and those of cannabis dependence span 1.3% to 2.2% (Agrawal&Lynskey, 2007; Stinson et al., 2005; von Sydow et al., 2001).

A central issue concerning the Diagnostic and Statistical Manual of Mental Disorders (4th ed.) (DSM-IV) CUD criteria is whether the criteria of abuse are distinct from those for dependence (American Psychiatric Association, 1980, 1987, 1994; Edwards et al., 1981). The current consensus is that a single factor captures most of the association between DSM-IV criteria within multiple substances criteria (Feingold & Rounsaville, 1995a, 1995b; Gillespie et al., 2007; Hartman et al., 2008; Langenbucher et al., 2004; Lynskey & Agrawal, 2007; Nelson et al., 1999; Teesson et al., 2002). These findings are reflected in DSM-5, which removed the abuse-dependence distinction along with the legal problems criterion while adding withdrawal and craving criteria. Most, but not all, of these studies (Baillie & Teesson, 2010) have been based on North American samples. Replication of these findings in other populations is warranted, and it is necessary to determine whether craving and withdrawal measure the same underlying dimension of liability to CUD.

A further issue is whether other models, notably latent class or factor mixture models (FMMs), fit the data better than factor analytic models. Under the latent class model, items correlate in the population because it consists of two or more subpopulations, which differ in the probability of response to at least one of the criteria. Within each class, item response probabilities are assumed to be independent, such that, for example, the probability of endorsing a tolerance item and an abuse item is simply the product of the two response probabilities. The FMM elaborates on latent class model by allowing for non-independence of item response probabilities within each class. The differences between these models have important implications for etiology, prevention, and treatment. For example, there might exist one class of people who are asymptomatic (low loadings on all criteria), another class whose members have high on abuse but not dependence criteria, and a third class whose members are high on both abuse and dependence criteria. Under both latent class and FMMs, it is possible to compute the probability that an individual belongs to a particular class, and this may be examined by validation against external criteria, such as treatment response or environmental risk factors. Recent model fitting has found little evidence to justify the use of latent class or FMMs when describing CUD (Baillie & Teesson, 2010; Gillespie et al., 2011b), but more research is needed. Muthen (2006) did find that FMMs fit

DSM alcohol use disorder (AUD) criteria in a selected sample better than conventional factor or latent class models, but his study was not population-based as it was performed only on the subset of respondents who endorsed criteria, and it was not a study of cannabis.

To identify the best representation of the population distribution of liability to DSM-V CUD, we apply latent common factor, latent class, and FMMs to data from a population-based sample of young adult Australians.

Methods

Participants

Data are from a population-based sample of young adult Australian twins and their non-twin siblings who are part of the ongoing Brisbane Longitudinal Twin Study (BLTS) at the Queensland Institute of Medical Research (QIMR). Described in detail elsewhere (Gillespie et al., 2012; Wright & Martin, 2004), the BLTS began in 1992 when twins and their family members were recruited in the greater Brisbane area, mainly through schools, but also via media appeals and by word of mouth as part of an ongoing, multi-wave study examining the development of moles at ages 12 and 14, cognition at age 16, and psychiatric diagnoses, brain imaging, and lifestyle and behavioral assessments in their early twenties. Data for the current analyses were collected between 2009 and 2011 as part of an ongoing US National Institutes of Health/National Institute of Drug Abuse (NIH/NIDA) project to study the genetic and environmental pathways to cannabis use, abuse, and dependence. Ascertainment began with adult twins and non-twin singleton siblings from the BLTS sample in order to obtain data from individuals who had passed through the age of maximum risk for the onset of cannabis use (typically 16–18 years) and cannabis-related problems. Response rates across the BLTS projects since 1992 range from 73% to 85% (Gillespie et al., 2012). To date, complete data were obtained from 626 twins (367 (58.6%) females and 259 males), aged 20 to 38 years.

Measures and Reliability

A computer-assisted telephone interview (CATI) protocol was used to obtain demographic and background data, together with DSM-IV criteria for cannabis (marijuana, hashish, tetrahydrocannabinol [THC], or ganja) abuse, dependence, craving, and withdrawal. The cannabis assessment began with basic screening criteria, initiation, and frequency of use measures. Following screening for ‘Have you ever used marijuana?’ (Yes/No), only subjects who endorsed either ‘Have you used marijuana six or more times in your life?’ or ‘Have you ever used marijuana 11 or more times in a month?’ were asked the abuse, dependence, withdrawal, and craving criteria.

Following previous analyses, which showed that including screening criteria in the analyses is effective (Gillespie et al., 2011b, 2012), these criteria were summed and recoded onto a 3-point ordinal ‘stem’ item, which was included in all analyses. The following coding system was implemented: 0 = never tried or used for less than six times in lifetime; 1 = tried and used for six or more times in lifetime; or 2 = tried and used for 11 or more times in a month. When the stem is coded as 0, all criteria are coded as missing, rather than 0, because there is

a non-zero probability that subjects would develop the criteria if they initiated cannabis use. Our rationale for including this stem item was that joint analysis of the stem and cannabis use symptoms produces asymptotically unbiased estimates of (1) the proportion of people in the population who would develop symptoms if they were to initiate cannabis (i.e., the symptom thresholds); (2) the correlation between liability to initiate use and the liability to endorse abuse and dependence criteria; and (3) correlations among the symptoms themselves (Gillespie et al., 2011b, 2012). This method therefore yields asymptotically unbiased estimates of factor loadings and other model parameters. Moreover, inclusion of stem score along with marginal maximum likelihood (MML) estimation produces parameter estimates that are valid for the entire population under study rather than only the subset selected to receive the symptom criteria (Kubarych et al., 2005).While the stem has three levels, the DSM criteria are binary.

In order to correspond to the 'failure to fulfill major role obligations' criterion, the two criteria, 'used often when doing something important' and 'stayed away from school or missed appointments because of use', were aggregated and scored positive if either of the symptoms was endorsed. Similarly, 'felt sick when cutting down or stopped use' and 'after not using cannabis, used to prevent sickness' were similarly aggregated to correspond to withdrawal.

Interviewer Training, Quality Control, and Informed Consent

All interviewers were selected from an experienced pool of QIMR staff who participated in a 2-week training session consisting of didactic instruction and supervised practice interviews. All interviewers conducted at least three interviews with community volunteer subjects under the supervision of a faculty trainer or senior staff member. Following consent, the CATI interviews were recorded for editing and quality control. For quality control and to prevent interviewer drift, 5% of the interviews were re-entered by an independent editor listening to the recorded interview on a continuing basis throughout the project. Informed consent was obtained from all subjects. Ethics approvals were obtained from the Human Research Ethics Committee at the QIMR and the Institute Review Panel at Virginia Commonwealth University.

Statistical Analyses

We fit latent factor, latent class, and FMMs to the cannabis use criteria and stem item data. Latent factor analysis (LFA; Spearman, 1904) accounts for covariation among observed indicators in terms of a reduced number of latent factors. In contrast, latent class analysis (LCA) assumes that correlations between symptoms arise because populations consist of subgroups that differ in their means or variances. Although LCA may be useful for defining and validating psychiatric phenotypes (Leoutsakos et al., 2010), minor differences between classes can make it difficult to distinguish one class from another. Further (Lazarsfeld & Henry, 1968), individuals within a class are considered to be homogeneous and are not distinguishable from one another (Muthen, 2006). FMMs represent a hybrid of the two methods (Dolan & Maas, 1998; Everitt, 1988; Jedidi et al., 1997; McLachlan&Peel, 2000;Muthen, 2006;Muthen&Shedden, 1999; Yung, 1997). By allowing individuals in each latent class to also vary along continuous dimensions (factors) of observed criteria, FMMs

can identify both subpopulations of similar individuals and quantify individual differences among those individuals. Factor models and FMMs impose an underlying parametric model on the data, LCAs do not.

For all three modeling approaches we used MML raw ordinal data analysis in the Mx software package (Neale et al., 2006). MML (Bock & Aitken, 1981) estimates model parameters by computing the joint likelihood of the latent factor(s) and the observed data. This is accomplished by integrating over the latent factor distribution using the 10-point Gauss–Hermite quadrature (Neale et al., 2006). For each quadrature point, the product of the quadrature weight and the conditional likelihood of the vector of criteria data is computed, and these products are summed. This approach is computationally efficient because the criteria are independent when conditioned on the factor.

While factor models are typically easy to estimate, latent class and FMMs are more prone to local solutions and estimation problems (Goodman, 1974; Hipp & Bauer, 2006). All models were fit repeatedly using different sets of starting values to verify that a global minimum for each model was obtained. Models were considered to have converged on the global solution when the maximum likelihood (ML) value (minimum −2 log likelihood value) was reached for multiple times with different initial parameter values.

Choice of Best Fitting, Most Parsimonious, and Most Interpretable Model

When we compare different factor models, such as the one-and two-factor models, in this analysis, for example, the difference between their likelihoods is asymptotically distributed as a chi-square, so we can use a likelihood difference test (Steiger, 1985). When we compare factor models with latent class or mixture models, however, the difference is not asymptotically distributed as a chi-square. Comparisons between these models require omnibus fit indices. These indices rely on ‘twice the negative log-likelihood’ (−2LL), which is an index of misfit, plus a parsimony adjustment to take into account model complexity. The Akaike Information Criterion (AIC; Akaike, 1987), Bayesian Information Criterion (BIC), and sample size-adjusted BIC (SABIC; Schwarz, 1978) are common and useful information criterion indices. When comparing models within fit indices, the model with the lowest −2LL value is indicative of the best fitting model whereas the lowest, the most negative AIC, BIC, and SABIC values are indicative of the most parsimonious fit. Parsimony is important because in ML estimation, log likelihoods will continue to decrease with additional model parameters, which can result in ‘over-fitting’. Indices of parsimony penalize models with the increasing number of parameters, thereby providing a balance between model complexity and model/data misfit, with AIC having the weakest penalty for additional parameters and BIC having the strongest penalty. Furthermore, the penalties for BIC and SABIC increase with sample size and the number of parameters, while AIC penalties depend only on the number of estimated parameters.

Simulations have shown that BIC can correctly discriminate between LCA and factor models (Markon & Krueger, 2004). Differences in BIC between any two LCA models can be interpreted as having corrected for expected effects of sampling variation, and are exponentially related to the posterior odds of one model versus another (Markon & Krueger, 2005). With sufficiently large samples, the BIC should correctly identify the best

approximating model even among non-nested alternative models (Barron & Cover, 1991; Markon & Krueger, 2005; Vereshchagin & Vitanyi, 2004). However, because our sample size of 626 is not large for this purpose, our results should be interpreted with caution, especially given the number of parameters estimated. Simulations (Nylund et al., 2007) have also shown that the BIC and SABIC (Schwarz, 1978) can outperform the AIC in complex structures in which symptoms have different endorsement probabilities for more than one latent class. Although parametric bootstrapping may provide a better discrimination between LCA and FMM models with different numbers of latent classes (Nylund et al., 2007), it is extremely demanding computationally and was not used. Selecting a final model should be based on statistical information, but among those that differ only slightly in their fit to the data, the model with the most interpretable parameter estimates is to be preferred. Subsequent prediction of, for example, clinical outcomes may further validate model selection.

Results

Criteria Endorsements

Endorsement frequencies for the stem and diagnostic criteria are shown in Table 1. By including age at interview as a covariate on criterion thresholds, we adjust for potential age-related cohort changes in symptom endorsement. For both males and females, the most commonly endorsed criterion was ‘trying to cut down or stop using’, although the endorsement rate was higher in males (25.5%) than in females (10.9%). The second most commonly endorsed criterion was ‘ending up taking a lot more than intended or planned’ for males and ‘having to use a lot more in order to get high’ for females. The least frequently endorsed criterion was ‘cannabis use resulting in legal problems or traffic accidents’. The second least commonly endorsed criterion for both males and females was ‘cannabis use causing physical problems or depression’.

Phenotypic Correlations and Eigenvalues

Table 2 displays the polychoric correlation matrix for the 12 criteria and the stem. The first four eigenvalues were 8.6, 1.6, 1.1, and 0.5; thus, although there were three eigenvalues greater than unity, the ratio of the first to second eigenvalue was very large (5.43).

Model Comparisons

Table 3 displays the fit statistics for the 1, 2, and 3 class models, 1 factor, two orthogonal factors, two correlated (oblique) factors and three orthogonal factors, and one factor/2 class and one factor/3 class models. The most parsimonious models are shown in bold. The one-factor model provided the best fit as judged by the BIC, whereas the correlated two-factor solution performed better in terms of the AIC and SABIC criteria. The correlation between the un-rotated factors for the two-factor oblique model was 0.51. The two-factor solution is the best fitting solution by the likelihood ratio (LR) test, as well, so we can safely ignore the lone BIC result.

We then used PROMAX rotation on the best fitting exploratory two-factor oblique solution in the software program SAS (2011). Factor loadings appear in Table 4. Based on the factor-

loading pattern, the first dimension can be interpreted as a general liability to CUD factor. It is defined by use and symptoms of abuse, dependence, withdrawal, and craving. Craving loaded very highly (0.92) on the general factor. The second dimension is an impairment factor defined by four symptoms with loadings of 0.40 and higher: unable to fulfill school or work obligations; use causing problems with other people; spending a lot of time obtaining cannabis, using and recovering from it; and cannabis use that causes interference with work, study, family, and friends. The correlation between the factors for the rotated solution was 0.44.

Discussion

This is the first study to compare the fit of latent factor, latent class, and FMMs to cannabis use, symptoms of cannabis abuse, dependence, and withdrawal, along with DSM-5-based craving in a population-based sample of young adult Australians. Even with the addition of the craving symptom, our results are commensurate with recent findings: latent factor models outperform both latent class and FMMs (Gillespie et al., 2011a, 2012). Although most of the observed aggregation between the physiological, behavioral, and cognitive components of CUD is best explained by a general liability to CUD factor, we found evidence for the second, clinically interpretable factor that captures important social and occupational impairment associated with frequent cannabis use. This second factor was moderately correlated with the general CUD factor.

Our results are not fully comparable with those reported previously. Among the reviewed papers that support the consensus of a single liability dimension for CUD (Baillie & Teesson, 2010; Compton et al., 2009; Feingold & Rounsaville, 1995a, 1995b; Gillespie et al., 2011a, 2012; Hartman et al., 2008; Langenbucher et al., 2004; Lynskey & Agrawal, 2007; Nelson et al., 1999; Teesson et al., 2002), only three provided comparative fit indices between competing factorial models or omnibus comparisons with latent class and FMMs (Baillie & Teesson, 2010; Gillespie et al., 2011a, 2012), while two reports fitted confirmatory factor models (Compton et al., 2009; Nelson et al., 1999). In some instances, fit indices to facilitate model comparisons were not provided (Hartman et al., 2008; Langenbucher et al., 2004). In others, there were only marginal differences between the one- and two-factor models (Baillie & Teesson, 2010; Teesson et al., 2002). In two instances, a two-factor solution actually provided a slightly better fit to the data (Feingold & Rounsaville, 1995a; Lynskey & Agrawal, 2007). To what extent the empirical support for unidimensional models for other illicit and licit substances also varies is beyond the scope of this paper. Nevertheless, it is important to acknowledge that model fitting is rarely equivocal and that the reports cited above based their conclusions on additional metrics: eigenvalues or eigenvalue ratios (Hartman et al., 2008; Langenbucher et al., 2004); low mean square residual values, and scant residual inter-item correlations (Langenbucher et al., 2004); poor interpretability of additional dimensions (Gillespie et al., 2011a); improvement in fit for the two-factor solution attributable to very large samples (Lynskey & Agrawal, 2007); or the very high observed inter-factor correlations (Lynskey &Agrawal, 2007; Teesson et al., 2002).

In contrast, our population-based sample was relatively small, and the inter-factor correlation following rotation was only moderate (r = 0.44). Moreover, the two-factor correlated solution was the most consistent solution across the fit indices. Although the first-to-second eigenvalue ratio suggests that the first dimension captures most of the covariance, the pattern of loadings on the second dimension is consistent with the observed statistics. Given the moderate inter-factor correlation, we speculate that there are individuals with high liability to CUD but low impairment, that is, resilience despite use. This makes clinical sense; despite frequent use and manifest signs of the more pharmacological aspects of cannabis addiction, including tolerance, withdrawal, and craving, a proportion of cannabis users can remain resilient in terms of normal functioning as defined by minimal social and occupational impairment. On the whole, our findings are in contrast to a growing consensus that a single factor can adequately account for the covariation among the cannabis criteria. Evidence for the second factor that includes clinically relevant features of addiction not captured by the general CUD factor is substantively plausible and etiologically relevant.

While overall heritability for general problematic cannabis use ranges from 51% in males to 59% in females (Verweij et al., 2010), estimates of genetic variance for the impairment symptoms have not been determined. A larger sample size is required. There is, however, evidence to support the role of genetic, psychosocial, and developmental components for correlated phenotypes such as resilience (Ahmed, 2012; Russo et al., 2012). To what extent the observed general CUD and impairment factors correspond to different genetic or environmental risks is unclear at this point. Twin studies have typically focused on either the genetics of use or at the syndrome levels of abuse or dependence, instead of at the item of symptom level. A recent multivariate genetic analysis of the criteria for DSM-IV alcohol dependence identified not one but three genetic liabilities indexing risk of (1) tolerance and heavy use; (2) loss of control with alcohol-associated social dysfunction; and (3) withdrawal and continued use despite problems (Kendler et al., 2012). These results are at odds with a single, coherent phenotypic factor structure (Beseler et al., 2010; Borges et al., 2010; Saha et al., 2006), but are consistent with rodent studies examining the genetic influences on a variety of alcohol-related traits: genetic contributions to each are either largely distinct or only weakly correlated (Crabbe et al., 1999, 2005). It therefore remains to be seen if similar complexity arises from cannabis use and symptoms of abuse, dependence, withdrawal, and craving.

This is also the first report to include craving in a combined analyses of cannabis use and DSM criteria. Reports examining the association between craving and symptoms of AUD have reported similar results (Bond et al., 2012; Cherpitel et al., 2010;Glockner-Rist et al., 2013;Hasin et al., 2012; Keyes et al., 2011). Based on the direct equivalence of the normal ogive item-response model (IRM) to factor-analysis of binary data (Takane & Leeuw, 1987), the symptom threshold and high factor loading for craving suggest that this symptom assesses lower levels of the liability to the general CUD factor with good to very good discrimination. Although a larger sample is required for a more definitive conclusion, the pattern of monozygotic (r_{mz} = 0.75) and dizygotic (r_{dz} = 0.38) polychoric twin pair correlations suggests that there is a high degree of familial aggregation in craving attributable to additive genetic risk factors. The legal problems criterion that was dropped in

DSM-5 was infrequently endorsed, implying a high IRM difficulty, but a high loading (0.92), implying that it discriminates extremely well at a very high level of liability to CUD. Legal problems may thus still be a useful criterion for identifying subjects at the highest level of liability to CUD.

The addition of craving also allowed for a comparison between the prevalence of DSM-IV abuse/dependence diagnoses with DSM-5 CUD. In fact, our analyses suggest that the prevalence of DSM-5 CUD will be slightly higher than that of DSM-IV cannabis abuse or dependence — 19.5% versus 16.9%, representing a modest increase of 15.4%. Removing craving did not alter this finding. By comparison, a recent analysis of data from the 2007 Australian National Survey of Mental Health and Wellbeing found that the prevalence of CUD decreased from6.2% according to DSM-IV criteria to 5.4% using DSM-5 criteria (Mewton et al., 2013). The trend observed with the current data is similar to the anticipated prevalence increase in DSM-5 AUD for North American samples (Agrawal et al., 2011; Edwards et al., 2013) but much lower than those in Australian samples (Mewton et al., 2011).

Limitations

Our findings must be interpreted in the context of at least four minor and two larger potential limitations. First, assessments were based on a single interview that necessarily included measurement error.

Second, lower endorsement rates of some criteria may have contributed to unstable parameter estimates.

Third, twin pair members were treated as independent observations. However, failure to take into account statistical non-independence is not expected to bias parameter estimates, but confidence intervals and fit indexes may be slightly underestimated. Based on our own published analyses, we speculated that non-independence of observations is rarely a problem when the group size is small. In the case of our twin data, the group size is at most two.

Fourth, model identification relied on the assumption that the cannabis stem(0 = never tried or used less than six times, 1 = tried and used six times or more, 2 = tried and used 11 or more times)was one-dimensional. We tested this assumption using monozygotic twin pairs and have found no evidence for its violation (Gillespie et al., 2007).

Fifth, only subjects who met a minimal threshold of cannabis use were administered the criteria. Consequently, our sample included a relatively large amount of ‘missing’ symptom data. Fortunately, the advantage of including an ordinal stem based on initiation (and use) in the analyses provides a means to predict whether or not symptoms are missing. In other words, the ordinal stem effectively corrects for the fact that we were missing data on abuse, dependence, withdrawal, and craving on subjects who denied ever using cannabis. Moreover, because ML estimates are robust to certain forms of ‘missingness’, it is reasonable to expect good recovery of the population values of the parameters. Including the ordinal stem also allows us to pose the following question: ‘How well does criterion×measure the latent trait or liability to develop symptoms of cannabis use disorder?’

Finally, although our latent class and FMMs can be a useful means of identifying and validating subpopulations of psychiatric phenotypes (Leoutsakos et al., 2010), the models are still prone to estimation problems. Latent class depends critically on the stringent assumption that local dependence holds, in that there is no residual covariance within a particular class and that whatever variance remains is due to measurement error. This can lead to the creation of classes with very minor differences, making it difficult to distinguish classes from one another (Gillespie & Neale, 2006). Distributional abnormalities can generate artifactual latent classes, which are typically substantively uninterpretable (Bauer & Curran, 2004). One possible explanation for the mixture models fitting less well than dimensional models in this application is that the interviews for substance abuse were not designed to both classify people into groups and measure individuals' liability (Clark et al., 2014); this is a matter for future research.

More generally, estimation problems can arise because of convergence on local rather than global solutions, thereby making it difficult to distinguish between models based on a single optimization. Our solution to this problem was to estimate each model multiple times using a range of possible starting or initial values for each estimated parameter and retaining the best-fitting solution from the entire set of estimated models (Goodman, 1974; Hipp & Bauer, 2006). While there is no requirement that the same solution be reached from multiple sets of starting values, greater numbers of convergences on the same solution increase confidence that a global rather than local solution has been found.

Notwithstanding the above limitations, when compared with factor mixture and latent class models, factor models provided a more parsimonious fit to the data. When conditioned on initiation and cannabis use, the association between the symptoms of cannabis abuse, dependence, withdrawal, and craving can be best explained by two correlated latent factors: a general risk factor to CUD; along with a factor defined by clinically relevant features assessing social and occupational impairment related to frequent cannabis use. Secondary analyses revealed that there is a modest increase in the prevalence of DSM-5 CUD compared with DSM-IV cannabis abuse or dependence.

Acknowledgments

Nathan Gillespie was supported by the United States National Institute on Drug Abuse (NIDA) K99R00 award (R00DA023549). Michael Neale was supported by NIDA grants DA-18673 and DA-21169. Ian Hickie was supported by an Australian National Health and Medical Research Council Australia Fellowship (No. 464914). We also acknowledge and thank the following QIMR project staff: Soad Hancock as project coordinator; David Smyth for IT; Lenore Sullivan as research editor; and our research interviewers Pieta-Marie Shertock and Jill Wood. We thank the twins and their siblings for their willing cooperation.

References

Agrawal A, Heath AC, Lynskey MT. DSM-IV to DSM-5: The impact of proposed revisions on diagnosis of alcohol use disorders. Addiction. 2011; 106:1935–1943. [PubMed: 21631621]

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]

Ahmed SH. The science of making drug-addicted animals. Neuroscience. 2012; 211:107–125. [PubMed: 21864653]

Akaike H. Factor analysis and AIC. Psychometrika. 1987; 52:317–332.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 3rd ed.. Washington, DC: Author; 1980.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-III-R. 3rd Revised ed.. Washington, DC: Author; 1987.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV. 4th ed.. Washington, DC: Author; 1994.

Baillie AJ, Teesson M. Continuous, categorical and mixture models of DSM-IV alcohol and cannabis use disorders in the Australian community. Addiction. 2010; 105:1246–1253. [PubMed: 20491729]

Barron AR, Cover TM. Minimum complexity density estimation. IEEE Transactions on Information Theory. 1991; 1:1034–1054.

Bauer DJ, Curran PJ. The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. Psychological Methods. 2004; 9:3–29. [PubMed: 15053717]

Beseler CL, Taylor LA, Leeman RF. An item-response theory analysis of DSM-IV alcohol-use disorder criteria and 'binge' drinking in undergraduates. Journal of Studies on Alcohol and Drugs. 2010; 71:418–423. [PubMed: 20409436]

Bock RD, Aitken M. Marginal maximum likelihood estimation of item parameters: An application of the EM algorithm. Psychometrika. 1981; 46:443–460.

Bond J, Ye Y, Cherpitel CJ, Borges G, Cremonte M, Moskalewicz J, Swiatkiewicz G. Scaling properties of the combined ICD-10 dependence and harms criteria and comparisons with DSM-5 alcohol use disorder criteria among patients in the emergency department. Journal of Studies on Alcohol and Drugs. 2012; 73(2):328–336. [PubMed: 22333341]

Borges G, Ye Y, Bond J, Cherpitel CJ, Cremonte M, Moskalewicz J, Rubio-Stipec M. The dimensionality of alcohol use disorders and alcohol consumption in a cross-national perspective. Addiction. 2010; 105:240–254. [PubMed: 20078482]

Cherpitel CJ, Borges G, Ye Y, Bond J, Cremonte M, Moskalewicz J, Swiatkiewicz G. Performance of a craving criterion in DSM alcohol use disorders. Journal of Studies on Alcohol and Drugs. 2010; 71(5):674–684. [PubMed: 20731972]

Clark S, Muthen BO, Kaprio J, D'Onofrio B, Vike R, Rose R. Models and strategies for factor mixture analysis: An example concerning the structure underlying psychological disorders. Structural Equation Modeling. 2014; 20:681–703.

Compton WM, Saha TD, Conway KP, Grant BF. The role of cannabis use within a dimensional approach to cannabis use disorders. Drug and Alcohol Dependence. 2009; 100:221–227. [PubMed: 19062204]

Crabbe JC, Metten P, Cameron AJ, Wahlsten D. An analysis of the genetics of alcohol intoxication in inbred mice. Neuroscience & Biobehavioral Reviews. 2005; 28:785–802. [PubMed: 15642621]

Crabbe JC, Phillips TJ, Buck KJ, Cunningham CL, Belknap JK. Identifying genes for alcohol and drug sensitivity: Recent progress and future directions. Trends in Neurosciences. 1999; 22:173–179. [PubMed: 10203855]

Dennis M, Babor TF, Roebuck MC, Donaldson J. Changing the focus: The case for recognizing and treating cannabis use disorders. Addiction. 2002; 97(Suppl. 1):4–15. [PubMed: 12460125]

Dolan CV, Maas HLJvd. 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; 59:225–242. [PubMed: 6972816]

Edwards AC, Gillespie NA, Aggen SH, Kendler KS. Assessment of a modified DSM-5 diagnosis of alcohol use disorder in a genetically informative population. Alcoholism, Clinical and Experimental Research. 2013; 37:443–451.

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; 90:1661–1669. [PubMed: 8555957]

Gillespie NA, Henders AK, Davenport TA, Hermens DF, Wright MJ, Martin NG, Hickie IB. The Brisbane Longitudinal Twin Study: Pathways to cannabis use, abuse, and dependence project-current status, preliminary results, and future directions. Twin Research and Human Genetics. 2012; 16(1):21–33. [PubMed: 23187020]

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. 2011a; 118:166–172. [PubMed: 21507586]

Gillespie NA, Kendler KS, Neale MC. Psychometric modeling of initiation and use and the symptoms of cannabis abuse, dependence and withdrawl in a sample of male and female twins. Drug and Alcohol Dependence. 2011b; 118:166–172. [PubMed: 21507586]

Gillespie NA, Neale MC. A finite mixture model for genotype and environment interactions: Detecting latent population heterogeneity. Twin Research and Human Genetics. 2006; 9:412–423. [PubMed: 16790151]

Gillespie NA, Neale MC, Legrand LN, Iacono WG, McGue M. 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. Psychology of Addictive Behaviors. 2012; 26:68–77. [PubMed: 22082343]

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]

Glockner-Rist A, Lemenager T, Mann K. Reward and relief craving tendencies in patients with alcohol use disorders: Results from the PREDICT study. Addictive Behaviors. 2013; 38:1532–1540. [PubMed: 23148916]

Goodman LA. Exploratory latent structure analysis using both identifiable and unidentifiable models. Biometrika. 1974; 61:215–231.

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. Toronto, Canada: Addiction Research Foundation; 1999. p. 69-125.

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; 47:165–173. [PubMed: 18176333]

Hasin DS, Fenton MC, Beseler C, Park JY, Wall MM. Analyses related to the development of DSM-5 criteria for substance use related disorders: 2. Proposed DSM-5 criteria for alcohol, cannabis, cocaine and heroin disorders in 663 substance abuse patients. Drug and Alcohol Dependence. 2012; 122:28–37. [PubMed: 21963333]

Hipp JR, Bauer DJ. Local solutions in the estimation of growth mixture models. Psychological Methods. 2006; 11:36–53. [PubMed: 16594766]

Jedidi K, Jagpal HS, Desarbo WS. Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. Marketing Science. 1997; 16:39–59.

Kendler KS, Aggen SH, Prescott CA, Crabbe J, Neale MC. Evidence for multiple genetic factors underlying the DSM-IV criteria for alcohol dependence. Molecular Psychiatry. 2012; 17:1306–1315. [PubMed: 22105626]

Keyes KM, Krueger RF, Grant BF, Hasin DS. Alcohol craving and the dimensionality of alcohol disorders. Psychological Medicine. 2011; 41:629–640. [PubMed: 20459881]

Kubarych TS, Aggen SH, Hettema JM, Kendler KS, Neale MC. Endorsement frequencies and factor structure of DSM-III-R and DSM-IV generalized anxiety disorder symptoms in women: Implications for future research, classification, treatment and comorbidity. International Journal of Methods in Psychiatric Research. 2005; 14(2):69–81. [PubMed: 16175876]

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]

Lazarsfeld, PF.; Henry, NW. Latent structure analysis. Boston, MA: Houghton Mifflin; 1968.

Leoutsakos JM, Zandi PP, Bandeen-Roche K, Lyketsos CG. Searching for valid psychiatric phenotypes: Discrete latent variable models. International Journal of Methods in Psychiatric Research. 2010; 19:63–73. [PubMed: 20187060]

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]

Markon, KE.; Krueger, RF. A modeling approach to distinguishing between discrete and continuous forms of psychopathology; Paper presented at the 19th Annual Meeting of the Society for Research in Psychopathology; St Louis, MO. 2004 Oct.

Markon KE, Krueger RF. Categorical and continuous models of liability to externalizing disorders: A direct comparison in NESARC. Archives of General Psychiatry. 2005; 62:1352–1359. [PubMed: 16330723]

McLachlan, GJ.; Peel, D. Finite mixture models. New York, NY: Wiley; 2000.

Mewton L, Slade T, McBride O, Grove R, Teesson M. An evaluation of the proposed DSM-5 alcohol use disorder criteria using Australian national data. Addiction. 2011; 106:941–950. [PubMed: 21205055]

Mewton L, Slade T, Teeson M. An evaluation of the proposed DSM-5 cannabis use disorder criteria using Australian national survey data. Journal of Studies on Alcohol and Drugs. 2013; 74:614–621. [PubMed: 23739026]

Muthen B. Should substance use disorders be considered as categorical or dimensional? Addiction. 2006; 101(Suppl. 1):6–16. [PubMed: 16930156]

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; 31:1010–1034. [PubMed: 16723188]

Neale, MC.; Boker, SM.; Xie, G.; Maes, HH. Mx: Statistical modeling. 7th ed.. Richmond, VA: Department of Psychiatry, Medical College of Virginia; 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; 94:843–855. [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.

Russo SJ, Murrough JW, Han MH, Charney DS, Nestler EJ. Neurobiology of resilience. Nature Neuroscience. 2012; 15:1475–1484.

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]

SAS. SAS/STATR® 9.3 user's guide. Cary, NC: SAS Institute; 2011.

Schwarz G. Estimating the dimension of a model. Annals of Statistics. 1978; 6:461–464.

Spearman C. General intelligence, objectively determined and measured. American Journal of Psychology. 1904; 4:201–293.

Steiger J. On the multivariate asymptotic distribution of sequential chi-square tests. Psychometrika. 1985; 50:253–264.

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; 80:105–116. [PubMed: 16157233]

Takane Y, Leeuw JD. On the relationship between item response theory and factor analysis of discretized variables. Psychometrika. 1987; 52:393–408.

Teesson M, Lynskey M, Manor B, Baillie A. The structure of cannabis dependence in the community. Drug and Alcohol Dependence. 2002; 68:255–262. [PubMed: 12393220]

Vereshchagin NK, Vitanyi PMB. Kolmogorov's structure functions and model selection. IEEE Transactions on Information Theory. 2004; 50:3265–3290.

Verweij KJ, Zietsch BP, Lynskey MT, Medland SE, Neale MC, Martin NG, Vink JM. Genetic and environmental influences on cannabis use initiation and problematic use: A meta-analysis of twin studies. Addiction. 2010; 105:417–430. [PubMed: 20402985]

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; 64:347–361. [PubMed: 11672949]

Wright MJ, Martin NG. The Brisbane Adolescent Twin Study: Outline of study methods and research projects. Australian Journal of Psychology. 2004; 56:65–78.

Yung YF. Finite mixtures in confirmatory factor-analysis models. Psychometrika. 1997; 62:297–330.

TABLE 1

Rates of Endorsement for the Ordinal Stem and Diagnostic Criteria

	Males	Females
Ordinal stem[a]	44.0% (*n* = 114)	25.1% (*n* = 92)
Failure to fulfill major role obligation at work, school, or home[b]	13.9% (*n* = 36)	4.1% (*n* = 15)
Ever use it in a situation in which it might have been physically dangerous?	17.8% (*n* = 46)	5.4% (*n* = 20)
Have legal problems or traffic accidents because you were using marijuana?	4.6% (*n* = 12)	0.3% (*n* = 1)
Using it causes problems with other people?	10.8% (*n* = 28)	4.6% (*n* = 17)
Use a lot more in order to get high or feel its effects compared with when you first started?	16.6% (*n* = 43)	9.0% (33)
Withdrawal symptoms[c]	15.4% (*n* = 40)	8.17% (*n* = 30)
Ended up taking a lot more than you intended or planned?	21.2% (*n* = 55)	8.7% (*n* = 32)
Desire or attempt to cut down?	25.5% (*n* = 66)	10.9% (*n* = 40)
Spend a lot of time using it, recovering from using it, or doing whatever you had to do to get it?	13.1% (*n* = 34)	4.6% (*n* = 17)
Take it so often … instead of working, studying … or spending time with family and friends?	9.6% (*n* = 25)	4.1% (*n* = 15)
Using it causes you physical problems, or makes you depressed or very nervous?	8.1% (*n* = 21)	3.8% (*n* = 14)
Ever crave, desire, or have an urge for smoking marijuana?	18.5% (*n* = 48)	9.3% (*n* = 34)

Note: All items were prefaced with, ‘During this time when you used cannabis the most did you … ?’

[a] 0 = Never tried or tried but never for more than six times in lifetime, 1 = tried and had used for more than six times in lifetime, 2 = tried and had used for 11+ times in a month; endorsement rates reflect percentage who endorsed 1 or 2 on the stem.

[b] Aggregate of ‘Use while doing something important like being at school or work or taking care of children?’ and ‘Stay away from work or miss appointments because you were using it?’

[c] Aggregate of ‘Did you ever have one or more of the withdrawal symptoms in the list?’ and ‘Use it to relieve, stop, or avoid getting sick or withdrawal symptoms?’

TABLE 2

FIML-Estimated and Age- and Sex-Adjusted Polychoric Correlation Matrix of CUD Stem, Abuse, and Dependence Criteria

	Stem	1	2	3	4	5	6	7	8	9	10	11	12
Ordinal stem[a]	1.00												
1. Failure to fulfill major role obligation at work school or home[b]	0.65	1.00											
2. Ever use it in a situation in which it might have been physically dangerous?	0.54	0.77	1.00										
3. Have legal problems or traffic accidents because you were using marijuana?	0.44	0.48	0.38	1.00									
4. Using it causes problems with other people?	0.49	0.76	0.66	0.40	1.00								
5. Use a lot more in order to get high or feel its effects compared with when you first started?	0.76	0.71	0.67	0.85	0.53	1.00							
6. Withdrawal symptoms[c]	0.61	0.68	0.61	0.87	0.73	0.84	1.00						
7. Ended up taking a lot more than you intended or planned?	0.58	0.62	0.65	0.58	0.66	0.74	0.72	1.00					
8. Desire or attempt to cut down? 1, 4, or 5	−0.06	0.16	0.30	0.39	0.40	0.34	0.41	0.47	1.00				
9. Spend a lot of time using, recovering from using, or doing whatever you had to do to get it?	0.78	0.81	0.64	0.55	0.82	0.78	0.74	0.71	0.16	1.00			
10. Take it so often … instead of working, studying … or spending time with family and friends?	0.80	0.79	0.69	0.50	0.77	0.79	0.67	0.79	0.17	0.96	1.00		
11. Using it causes you physical problems, or makes you depressed or very nervous?	0.16	0.59	0.56	0.50	0.79	0.49	0.62	0.66	0.66	0.54	0.57	1.00	
12. Ever crave, desire, or have an urge for smoking marijuana?	0.81	0.70	0.65	0.77	0.63	0.90	0.87	0.79	0.27	0.79	0.81	0.52	1.00

Note: All items were prefaced with ‘During this time when you used cannabis the most did you …?’

[a] 0 = Never tried or tried but never for more than six times in lifetime, 1 = tried and had used for more than six times in lifetime, 2 = tried and had used for 11+ times in a month.

[b] Aggregate of ‘Use while doing something important like being at school or work or taking care of children?’ and ‘Stay away from work or miss appointments because you were using it?’

[c] Aggregate of ‘Did you ever have one or more of the withdrawal symptoms in the list?’ and ‘Use it to relieve, stop, or avoid getting sick or withdrawal symptoms?’

TABLE 3

Comparison of Latent Factor, Latent Class, and Factor Mixture Models for Cannabis Symptoms

Models	−2LL	#Par.	AIC	BIC	SABIC
Latent class					
1-Class	3,799	40	−2,319	−7,950	−3,094
2-Class	2,986	56	−3,099	−8,304	−3,474
3-Class	2,868	73	−3,184	−8,309	−3,505
Latent factor					
1-Factor	2,932	53	−3,158	**−8,338**	−3,504
2-Factor orthogonal	2,883	65	−3,183	−8,324	−3,509
2-Factor oblique	2,878	66	**−3,186**	−8,323	**−3,510**
3-Factor orthogonal	2,877	76	−3,167	−8,291	−3,494
Factor mixture					
1-Factor, 2 classes	2,804	108	−3,178	−8,228	−3,480
1-Factor, 3 classes	2,761	162	−3,113	−8,076	−3,414

Note: −2LL = −2 × log likelihood, #Par. = number of estimated parameters, AIC = Akaike's Information Criteria, BIC = Bayesian Information Criterion, SABIC = Sample Size-Adjusted Bayesian Information Criteria.
All models included age and sex as covariates on the symptom and stem-item thresholds.
The best fitting model for AIC, BIC, and SABIC criterion are in bold.

TABLE 4

Factor Loadings for the Best Fitting Two-Factor Oblique Solution for Cannabis Symptoms and Stem Item (*N* = 626) Following PROMAX Rotation in SAS

	F1	F2
Ordinal stem[a]	**0.98**	−0.11
Failure to fulfill major role obligation at work school or home[b]	0.34	**0.48**
Ever used it in a situation in which it might have been physically dangerous?	**0.49**	0.30
Have legal problems or traffic accidents because you were using marijuana?	**0.94**	−0.23
Using it causes problems with other people?	**0.48**	**0.40**
Use a lot more in order to get high or feel its effects compared with when you first started?	**0.80**	0.11
Withdrawal symptoms[c]	**0.82**	0.05
Ended up taking a lot more than you intended or planned?	**0.82**	0.06
Desire or attempt to cut down?	**0.64**	−0.04
Spend a lot of time using it, recovering from using it, or doing whatever you had to do to get it?	0.29	**0.62**
Take it so often … instead of working, studying … or spending time with family and friends?	0.09	**0.80**
Using it causes you physical problems or makes you depressed or very nervous?	**0.65**	0.21
Ever crave, desire, or have an urge for smoking marijuana?	**0.92**	−0.02

Note: The correlation between the factors is $r = 0.51$.

[a] 0 = Never tried or tried but never for more than six times in lifetime, 1 = tried and had used for more than six times in lifetime, 2 = tried and had used for 11+ times in a month.

[b] Aggregate of ‘Use while doing something important like being at school or work, or taking care of children?’ and ‘Stay away from work or miss appointments because you were using it?’

[c] Aggregate of ‘Did you ever have one or more of the withdrawal symptoms in the list?’ and ‘Use it to relieve, stop, or avoid getting sick or withdrawal symptoms?’