Psychometric modeling of cannabis initiation and use and the symptoms of cannabis abuse, dependence and withdrawal in a sample of male and female twins

Nathan A. Gillespie\textsuperscript{a,b,*}, Kenneth S. Kendler\textsuperscript{a,c}, Michael C. Neale\textsuperscript{a,c}

\textsuperscript{a} Department of Psychiatry, Virginia Commonwealth University, United States
\textsuperscript{b} Queensland Institute of Medical Research, Brisbane, Australia
\textsuperscript{c} Department of Human Genetics, Virginia Commonwealth University, United States

\textbf{Article Info}

Article history:
Received 21 July 2010
Received in revised form 28 February 2011
Accepted 17 March 2011
Available online xxx

Keywords:
Cannabis
Abuse
Dependence
Withdrawal
Heterogeneity
Latent class
Latent factor
Factor mixture

\textbf{Abstract}

\textbf{Background:} Despite an emerging consensus that the DSM-IV diagnostic criteria for cannabis abuse and dependence are best represented by a single underlying liability, it remains unknown if latent class or hybrid models can better explain the data.

\textbf{Method:} Using structured interviews, 7316 adult male and female twins provided complete data on DSM-IV symptoms of cannabis abuse and dependence. Our aim was to derive a parsimonious, best-fitting cannabis use disorder (CUD) phenotype based on DSM-III-R/IV criteria by comparing an array of psychometric models (latent factor analysis, latent class analysis and factor mixture modeling) using full information maximum likelihood ordinal data methods in Mx.

\textbf{Results:} We found little evidence to support population heterogeneity since neither latent class nor hybrid factor mixture models provided a consistently good fit to the data. When conditioned on initiation and cannabis use, the endorsement patterns of the abuse, dependence and withdrawal criteria were best explained by two latent factors for males and females. The first was a general CUD factor for which genetic effects explained 53–54% of the variance. A less interpretable second factor included a mix of cross-loading dependence and withdrawal symptoms.

\textbf{Conclusions:} This is the first study to compare competing measurement models to derive an empirically determined CUD phenotype. Commensurate with proposed changes to substance use disorders in the DSM-V, our results support an emerging consensus that a single CUD latent factor can more optimally assess the risk or liability underpinning correlated measures of use, abuse, dependence and withdrawal criterion.

\section{1. Introduction}

Cannabis is the most widely used illicit drug in developed countries including the United States (Dennis et al., 2002; Hall et al., 1999). Population based estimates of lifetime cannabis use in the USA between 1990 and 2004 range from 41.2% to 55.9% (Agrawal et al., 2004; Degenhardt et al., 2008; Kendler et al., 2003, 2000; Samhsa, 1998; von Sydow et al., 2001). Estimates of the lifetime prevalence of abuse range from 5.5% to 8.4% whereas cannabis dependence ranges from 1.3% to 2.2% (Agrawal and Lynskey, 2007; Stinson et al., 2005; von Sydow et al., 2001). A central issue in interpreting the available epidemiologic data on cannabis use has been the assumption that the DSM-IV symptoms of cannabis abuse and dependence captured two distinct, categorical outcomes (American Psychiatric Association, 1980, 1987, 1994; Edwards et al., 1981).

Recent attempts to understand why covariation between the symptoms of cannabis abuse and the dependence exist have steadily challenged this assumption and the DSM categorical model. It is now commonly accepted that the covariance arises because of a general factor that influences all cannabis symptoms (Feingold and Rounsaville, 1995a,b; Gillespie et al., 2007; Hartman et al., 2008; Langenbucher et al., 2004; Lynskey and Agrawal, 2007; Nelson et al., 1999; Teesson et al., 2002). By relying on clinical and population-based samples these reports have shown how the physiological, behavioral and cognitive components of cannabis abuse and dependence, including withdrawal, are consistent with a single psychometric factor model.
Overall, the factor model approach (Spearman, 1904) has been useful for developing and validating constructs when investigating the use and abuse of either a single drug or variety of substances (Harford and Muthen, 2001; Hasin et al., 1997; Lennox et al., 1996; Nelson et al., 1999). Indeed, similar results have been reported for alcohol, cocaine, opiates and sedatives (Feingold and Rounsaville, 1985a;b; Langenbucher et al., 2004; Nelson et al., 1999). Within the framework of item response theory, it can also provide a powerful means of calibrating correlated criteria in terms of item difficulty and discrimination, and for developing phenotypic measures with the most informative and robust items (Kirisci et al., 2009). However, in the current context of cannabis users, the approach depends critically on the population being homogeneous, i.e., the same model applies to all individuals.

In the absence of model comparisons we cannot a priori rule out alternative, better fitting approaches particularly when plausibly, there exist population subgroups of cannabis users. For instance, there have been significant cohort differences in the prevalence of cannabis use in the US between 1934 and 1974 (Kandel et al., 2005), with higher rates of use, abuse and dependence reported among males (Agrawal and Lynskey, 2004; Agrawal et al., 2004; Kandel et al., 2005; von Sydow et al., 2001). These cohort differences may be qualitative, rather than quantitative in origin (Degenhardt et al., 2000; Johnson and Gerstein, 1998; Muller and Gmel, 2002). The trajectory from cannabis initiation to use, abuse and dependence (CUD) can also vary between subjects (phenotypic heterogeneity) or the same disorder may result from different sets of risk factors (etiological heterogeneity) (Ellickson et al., 2004; Kandel and Chen, 2000; Windle and Wiesner, 2004).

Latent class analyses (LCA) are appropriate for modeling heterogeneous data because they assume that correlations between symptoms arise not because of some unmeasured, latent factor but because populations consist of subgroups differing in their means or variances. Although LCA is relatively unstructured, insofar as entirely different response probability sets can exist for different classes, it has been used to identify severity spectrums (Lazarsfeld and Henry, 1968; Moustaki, 1996; Pickles and Angold, 2003; Waller and Meehl, 1997) for alcohol (Bucholz et al., 1996; Lynskey et al., 2005), cannabis use (Grant et al., 2006), and to study treatment response and evaluation (Grant et al., 2006). It is a useful method for defining and validating psychiatric phenotypes (Leoutsakos et al., 2010). However, a limitation of LCA is that very minor differences between classes can make it difficult to distinguish one from the next and because of the local independence assumption (Lazarsfeld and Henry, 1968) it is impossible, unlike LFA, to distinguish individuals within a class from one another, e.g., individual within-class differences in severity (Muthen, 2006).

Historically, the latent factor and class approaches have been applied independently and for different purposes when in reality each represents a competing hypothesis regarding the covariation of cannabis diagnostic criteria. An alternative is a factor mixture model (FMM) that circumvents the limitations of LFA and LCA by applying a useful bridge between the two methods (Dolan and Maas, 1998; Everitt, 1988; Jeddidi et al., 1997; McLachlan and Peel, 2000; Muthen, 2006; Muthen and Shedden, 1999; Yung, 1997). FMMs can be used to identify latent classes on the one hand, while assessing whether there exist individual differences along continuous dimensions within each of the classes. In principle, this hybrid modeling strategy may fit the data better and if so, provide a superior measurement and classification for complex phenotypes. To date FMMs have been applied to symptoms of alcohol and nicotine consumption and arguably provide a richer representation of the data when compared to the fit of LFA or LCA models (Muthen, 2006; Muthen and Asparouhov, 2006). FMMs can also be applied to longitudinal data to identify subgroups with different growth trajectories. For example, a growth FMM was applied to antisocial behavior and heavy drinking (Muthen and Muthen, 2000) as well as to assess intervention effects of reducing aggressive behavior (Muthen et al., 2002).

We argue that supposition ought to be replaced with empirical evidence wherever possible. In this case, it has been supposed that the factor model is correct, but it has not been put to empirical test. While one would not want to expend effort rejecting models that are implausible, this not the case for LCA or FMMs of cannabis use patterns because if population heterogeneity exists, then it could prove very useful in the context of prevention or treatment. Since we know of no published reports that have compared LFA, LCAs and FMMs fitted to cannabis use disorder (CUD) symptoms our aim is to run omnibus comparisons between these models using DSM-III-R and DSM-IV cannabis diagnostic criteria. This will enable us to derive a best fitting and parsimonious CUD phenotype from existing, widely used assessment items which can potentially improve our understanding of the etiology and the prevention and treatment of CUDs.

2. Methods

2.1. Participants

Described in detail elsewhere (Kandel and Prescott, 2006), male data came from a 2nd wave of interviews between 1994 and 1998. Subjects were eligible for participation if they were at least in the 1st wave if they were successfully matched to birth records, or their cotwin was; a member of a multiple birth with at least one male; Caucasian; and born between 1940 and 1974. Of 9417 eligible individuals for the 1st wave (1993–1996), 6814 (72.4%) completed the interview. At least one year later, we contacted male–male and opposite-sex male–female twin pairs who had completed the initial interview to schedule a second interview. The 2nd interview was completed by 5629 individuals (82.6%).

As part of another ongoing study using the Mid Atlantic Twin Registry, the female data came from a 4th wave of interviews (1995–1997) of same-sex female–female twin pairs. These twins were eligible for participation if one or both twins were successfully matched to birth records; a member of a multiple birth with at least one female; Caucasian, and born between 1933 and 1972. At the 4th wave, ~2400 twins were potentially eligible from whom complete data were collected from 1928 individuals (~80%). These data were augmented with females from the 2nd interview (1994–1996) described above.

Complete cannabis initiation data were available from 4013 male subjects ranging in age from 20 to 58 years (μ = 37.0 yrs, σ2 = 9.1 yrs). Unlike previous analyses, these data (Gillespie et al., 2007) included an additional 1602 males from opposite-sex and incomplete twin pairs. Complete cannabis initiation data were available from 3303 female subjects ranging in age from 21 to 63 years (μ = 36.5 yrs, σ2 = 8.5 yrs).

2.2. Measures and reliability

All data sets included identical measures of cannabis lifetime abuse and dependence which were assessed using an adaptation of the Structured Clinical Interview for DSM-III-R-Patient Version (Spitzer et al., 1987) which was structured to allow assignment of DSM-IV drug disorder diagnoses. The eleven symptoms shown in Table 1 assessed four abuse and seven dependence criteria on a three-point scale [Definite/Probable/No]. ‘Definite’ and ‘Probable’ were combined in order to analyse each symptom as a dichotomous outcome.

In order to correspond to the first DSM-IV abuse criterion, item six is an aggregate of “used often when doing something important” or “stayed away from school or missed appointments because of use.” Similarly, item 12 is an aggregate of “felt sick when cutting down or stopped use” or “after not using cannabis, used to prevent sickness” to correspond with the DSM-III-R withdrawal.

Following screening for having ever taken cannabis (yes/no), administration of the drug abuse and dependence symptoms was contingent upon responses to two binary (yes/no) items measuring cannabis use. Subjects who endorsed using cannabis ‘6 or more times in a lifetime’ and not ‘11 times in a month’ were only asked abuse symptoms. However, if any abuse symptoms were endorsed then the dependence symptoms were also asked. Subjects who endorsed more frequent cannabis use ‘11 times in a month’ were automatically asked all symptoms. To determine whether they provided effective screening, the screening items were summed and recoded onto a three-point ordinal scale ‘stem item’ and included in the analyses. Our rationale for including the stem was that joint analysis of the stem and substance-use symptoms should produce asymptotically unbiased estimates of (i) the proportion of people in general the population who would develop symptoms if they were to initiate cannabis (i.e., the probes’ thresholds) and (ii) the correlation between liability to initiate and the liability to endorse abuse and dependence (represented by the factor loadings).
adjusted BIC (Schwarz, 1978) indices. The lowest, most negative
FMM models with different numbers of latent classes (Nylund et al., 2007) with parametric bootstrapping may provide a better discrimination between LCA and (Schwarz, 1978) can outperform the AIC in complex structures in which symptoms
of these conditional single item responses provides the overall likelihood for an for each item at a series of abscissae along the latent trait distribution. The product
1981) provides a fast means of calculating model parameters by integrating the (Neale et al., 2006b) to model the cannabis diagnostic criteria. MML (Bock and Aitken,
2.3. Statistical analyses
We used marginal maximum likelihood (MML) raw ordinal data analysis in Mx (Neale et al., 2006b) to model the cannabis diagnostic criteria. MML (Bock and Aitken, 1981) provides a fast means of calculating model parameters by integrating the latent factor distribution using 10-point quadrature (Neale et al., 2006a). 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. The product
of these conditional single item responses provides the overall likelihood for an observed vector of item scores. This product of one-dimensional integrals can be used because when conditioned on the factor, the items are independent.

2.4. Choice of best fitting and most parsimonious model

Latent factor, latent class, and factor mixture models were fitted to the ordinal data and compared using omnibus fit indices. In addition to inspecting the ‘minus twice log likelihood’ (−2LL) and the traditional Akaike Information Criterion (AIC) (Akaike, 1987) we used the Bayesian Information Criterion (BIC) and sample-size adjusted BIC (Schwarz, 1978) indices. 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 ‘overfitting’ 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.

Simulations have shown that BIC indices can correctly discriminate between LCA and LFA models (Markon and Krueger, 2004); the difference between any 2 BICs can be directly interpreted as having corrected for expected effects of sampling variation, and is exponentially related to the posterior odds of one model versus another (Markon and Krueger, 2005). Moreover, the BIC should correctly identify the best approximating model in large samples, even among non-nested alternatives (Barron and Cover, 1991; Markon and Krueger, 2005; Ververchagina and Vitanyi, 2004) thus providing a good omnibus comparison which is central to this paper’s aim.

Our decision to include the BIC and sBIC was also based on simulations (Nylund et al., 2007) that have shown how the 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. Although parametric bootstrapping may provide a better discrimination between LCA and FMM models with different numbers of latent classes (Nylund et al., 2007) it was not used because of its heavy computational burden.

3. Results

3.1. Item endorsements

Table 1 includes the item endorsements for each of the stem and diagnostic criteria. We formally modeled age at interview on item thresholds to adjust for potential age-related cohort changes in symptom endorsement. Among users, the two most commonly endorsed items for males and females are using cannabis in dangerous or hazardous situations and trying to cut down or stop using. The two least endorsed items are cannabis use resulting in legal problems or traffic accidents and withdrawal symptoms which include either feeling sick when cutting down or stopping use or using cannabis to prevent sickness.

Table 1

<table>
<thead>
<tr>
<th>Item endorsements for the stem and diagnostic criteria.</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Stem item</td>
<td>36.4% (N = 4013)</td>
<td>16.9% (N = 3245)</td>
</tr>
<tr>
<td>2. Used cannabis in dangerous situations</td>
<td>39.9% (N = 1460)</td>
<td>21.4% (N = 757)</td>
</tr>
<tr>
<td>3. Use caused legal problems or traffic accidents</td>
<td>4.4% (N = 1460)</td>
<td>1.7% (N = 757)</td>
</tr>
<tr>
<td>4. Use caused problems with family, friends, or people at work</td>
<td>15.2% (N = 1458)</td>
<td>11.3% (N = 755)</td>
</tr>
<tr>
<td>5. Use caused physical problems, made depressed or very nervous</td>
<td>13.0% (N = 1459)</td>
<td>10.9% (N = 755)</td>
</tr>
<tr>
<td>6. Failure to fulfill major role obligation at work school or home</td>
<td>28.2% (N = 1459)</td>
<td>18.6% (N = 757)</td>
</tr>
<tr>
<td>7. When started use ended up taking more than planned</td>
<td>26.0% (N = 976)</td>
<td>31.5% (N = 389)</td>
</tr>
<tr>
<td>8. Tried to cut down or stop using it</td>
<td>43.9% (N = 976)</td>
<td>44.4% (N = 390)</td>
</tr>
<tr>
<td>9. Spent time taking or using it, recovering from it, or doing whatever to get it</td>
<td>22.5% (N = 976)</td>
<td>18.3% (N = 389)</td>
</tr>
<tr>
<td>10. Used so often instead of work, hobbies or spending time with family or friends</td>
<td>12.7% (N = 975)</td>
<td>9.5% (N = 389)</td>
</tr>
<tr>
<td>11. Needed to use a lot more to get high or feel effects than when first used</td>
<td>31.3% (N = 976)</td>
<td>28.7% (N = 390)</td>
</tr>
<tr>
<td>12. Withdrawal symptoms a</td>
<td>6.3% (N = 976)</td>
<td>5.7% (N = 389)</td>
</tr>
</tbody>
</table>

a Stem coding: 0 = never tried or tried but never ≥11 times per month or ≥6 times lifetime, 1 = tried and used ≥11 times per month and ≥6 times lifetime. Represents proportion of subjects with stem ≥1

b Aggregation of used often when doing something important OR stayed away from school or missed appointments because of use.

c Aggregation of felt sick when cutting down or stopped use or after not using cannabis, used to prevent sickness.

3.2. Phenotypic correlations and eigenvalues

Polychoric correlations for the stem and diagnostic criteria appear in Table 2. Except for item 8 ‘Tried to cut down…stop using it’ most of the correlations were moderate to high. The first three eigenvalues for males and females respectively were 7.19, 1.18, 0.75 and 7.32, 1.08, 0.98, suggesting either one or two factor dimensional structure to the data.

3.3. Model comparisons

Fit statistics for males and females with the two best models highlighted under each fit index appear in Table 3. For both sexes, a 1-factor 3-class FMM solution provided a good fit as judged by the AIC. However, it was the 2-factor LFA solution that consistently performed better across all three fit indices. Varimax-rotated factor loadings based on 2-factor LFA appear in Table 4. For males and females this included a general use, abuse and dependence dimension. Only one item “Tried to cut down or stop using it” did not load onto this factor.

All of the dependence items cross-loaded onto a second factor for males. However, for females, the second factor was less interpretable; it included two dependence “Tried to cut down or stop using it” and “When started use ended up taking more than planned” and one abuse item “Use caused problems with family, friends, or people at work” cross-loaded onto this dimension.

4. Discussion

We found little evidence for population heterogeneity among cannabis users. Instead, our findings were commensurate with the consensus for a single factor that captures most of the association between the cannabis symptoms (Feingold and Rounsaville, 1995a,b; Hartman et al., 2008; Langenbucher et al., 2004; Lynskey and Agrawal, 2007; Nelson et al., 1999; Teesson et al., 2002). This consensus is anticipated in the forthcoming DSM-V that proposes to remove the current abuse-dependence distinction as well as revive the withdrawal criteria. With the exception of one item, all of the abuse and dependence symptoms as well as the stem loaded onto a single, general cannabis use disorder (CUD) factor. Although any
### Table 2

Age adjusted phenotypic polychoric correlations for the male (below diagonal) and female (above diagonal) stem and diagnostic criteria.

<table>
<thead>
<tr>
<th></th>
<th>1. Stem item</th>
<th>2. Used cannabis in dangerous situations</th>
<th>3. Use caused legal problems or traffic accidents</th>
<th>4. Used cannabis when people at work or home</th>
<th>5. Use caused physical problems, made depressed or very nervous</th>
<th>6. Failure to fulfill major role obligation at work school or home</th>
<th>7. When started use ended up taking more than planned</th>
<th>8. Tried to cut down or stop using it</th>
<th>9. Spent time taking or using it, recovering from it, or doing whatever to get it</th>
<th>10. High or feel effects than when first used</th>
<th>11. Withdrawal symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Stem item</td>
<td><strong>1.00</strong></td>
<td>0.62</td>
<td>0.28</td>
<td>0.59</td>
<td>0.44</td>
<td>0.43</td>
<td>0.44</td>
<td>0.71</td>
<td>0.58</td>
<td>0.58</td>
<td>0.52</td>
</tr>
<tr>
<td>2. Used cannabis in dangerous situations</td>
<td>0.62</td>
<td><strong>1.00</strong></td>
<td>0.52</td>
<td>0.59</td>
<td>0.44</td>
<td>0.43</td>
<td>0.44</td>
<td>0.71</td>
<td>0.58</td>
<td>0.58</td>
<td>0.52</td>
</tr>
<tr>
<td>3. Use caused legal problems or traffic accidents</td>
<td>0.28</td>
<td>0.52</td>
<td><strong>1.00</strong></td>
<td>0.40</td>
<td>0.32</td>
<td>0.32</td>
<td>0.32</td>
<td>0.51</td>
<td>0.41</td>
<td>0.41</td>
<td>0.35</td>
</tr>
<tr>
<td>4. Used cannabis when people at work or home</td>
<td>0.59</td>
<td>0.59</td>
<td>0.40</td>
<td><strong>1.00</strong></td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.50</td>
<td>0.44</td>
<td>0.44</td>
<td>0.47</td>
</tr>
<tr>
<td>5. Use caused physical problems, made depressed or very nervous</td>
<td>0.44</td>
<td>0.43</td>
<td>0.32</td>
<td>0.67</td>
<td><strong>1.00</strong></td>
<td>0.71</td>
<td>0.71</td>
<td>0.51</td>
<td>0.47</td>
<td>0.47</td>
<td>0.47</td>
</tr>
<tr>
<td>6. Failure to fulfill major role obligation at work school or home</td>
<td>0.43</td>
<td>0.44</td>
<td>0.32</td>
<td>0.67</td>
<td>0.71</td>
<td><strong>1.00</strong></td>
<td>0.71</td>
<td>0.51</td>
<td>0.47</td>
<td>0.47</td>
<td>0.47</td>
</tr>
<tr>
<td>7. When started use ended up taking more than planned</td>
<td>0.44</td>
<td>0.44</td>
<td>0.32</td>
<td>0.71</td>
<td>0.71</td>
<td>0.71</td>
<td><strong>1.00</strong></td>
<td>0.51</td>
<td>0.47</td>
<td>0.47</td>
<td>0.47</td>
</tr>
<tr>
<td>8. Tried to cut down or stop using it</td>
<td>0.71</td>
<td>0.71</td>
<td>0.51</td>
<td>0.51</td>
<td>0.51</td>
<td>0.51</td>
<td>0.51</td>
<td><strong>1.00</strong></td>
<td>0.44</td>
<td>0.44</td>
<td>0.44</td>
</tr>
<tr>
<td>9. Spent time taking or using it, recovering from it, or doing whatever to get it</td>
<td>0.58</td>
<td>0.58</td>
<td>0.47</td>
<td>0.47</td>
<td>0.47</td>
<td>0.47</td>
<td>0.47</td>
<td>0.44</td>
<td><strong>1.00</strong></td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>10. High or feel effects than when first used</td>
<td>0.58</td>
<td>0.58</td>
<td>0.47</td>
<td>0.47</td>
<td>0.47</td>
<td>0.47</td>
<td>0.47</td>
<td>0.44</td>
<td>0.67</td>
<td><strong>1.00</strong></td>
<td>0.59</td>
</tr>
<tr>
<td>11. Withdrawal symptoms</td>
<td>0.52</td>
<td>0.52</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
<td>0.44</td>
<td>0.67</td>
<td>0.59</td>
<td><strong>1.00</strong></td>
</tr>
</tbody>
</table>

**Notes:**
- 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 multifactorial and that the joint distribution of this scale with tailing scales underlying other ordinal items is two normal.
- Aggregation of “Used so often instead of work, hobbies or spending time with family or friends”.
- Aggregation of “Used when doing something important”. If the subject used cannabis, used to prevent sickness.
- Aggregation of “Stayed away from school or missed appointments because of use.”
- Aggregation of “Felt sick when cutting down or stopped use OR after not using cannabis, used to prevent sickness.”

There are notable differences between the current and previous findings. Previously, we reported (Gillespie et al., 2007) a single factor solution for cannabis use, abuse and dependence. This result was based on a smaller subset of the male data whereas the current analyses were augmented with males from opposite-sex and unmatched twin pairs. We also aggregated “used often when doing something important” and “stayed away from school or missed appointments because of use” into a single measure to correspond to the first DSM-IV abuse criterion. Likewise, two withdrawal items “felt sick when cutting down or stopped use” and “after not using cannabis, used to prevent sickness” were aggregated to correspond to DSM-III-R withdrawal.

Another difference with previous findings concerns the size of the factor loading particularly for the “Use caused legal problems or traffic accidents” item. This raises questions about the item’s usefulness as discriminating marker for CUD. Although four papers based on varying sample size and measurement criteria (12-month versus lifetime) have reported moderate to high loadings (Agrawal, 2007; Lynskey and Agrawal, 2007; Nelson et al., 1999), two papers based on a small Australian sample of cannabis users (Teesson et al., 2002) and the US population based sample (National Epidemiologic Survey on Alcohol and Related Conditions (NESARC)) (Compton et al., 2009) found that legal problems loaded poorly onto the single factor solutions. This discrepancy in factor loadings indicates that the item’s utility in terms of discrimination (see Takane and Leeuw, 1987) varies considerably. However, if we disregard the small or clinical samples (Feingold and Rounsaville, 1995a; Nelson et al., 1999; Teesson et al., 2002) and focus on the two large population based NESARC studies, then it is possible that the discrepancy is attributable in part to the 12-month (Compton et al., 2009) versus lifetime criteria (Lynskey and Agrawal, 2007). Agrawal and Lynskey (2007) in a re-analysis of the NESARC data using a life-time criteria found that although the item was endorsed at higher levels of risk among female users it was less discriminating by virtue of its relatively lower factor loading. In the absence of formal tests of sex invariance it is unclear if there was a significant sex difference in terms of item discrimination. In our sample, the lower factor loadings for females (0.50 versus 0.68) does suggest that legal problems may in fact discriminate better among male users.
Table 3
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 analyses also include oblique (r) solutions. Best fitting model is based on consistent performance across three fit indices. Bolding indicates top two under each index.

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>LFA – 1 dimension</td>
<td>18097.07, 17127, −16156.93, −62005.35, −34794.35</td>
<td>8705.13, 9325, −9944.87, −33343.15, −18528.34</td>
</tr>
<tr>
<td>LFA – 2 dimensions</td>
<td>17942.46, 17116, −16289.54, −62037.02, −34843.49</td>
<td>8659.90, 9314, −9968.10, −33321.29, −18523.97</td>
</tr>
<tr>
<td>LFA – 2 dimensions r</td>
<td>17942.46, 17115, −16287.54, −62032.87, −34840.94</td>
<td>8659.90, 9313, −9966.10, −33317.25, −18521.51</td>
</tr>
<tr>
<td>LFA – 3 dimensions</td>
<td>17964.42, 17087, −16209.58, −61905.73, −34758.28</td>
<td>19671.17, 17140, −14608.83, −61272.23, −34040.58</td>
</tr>
<tr>
<td>LCA – 1 class</td>
<td>18147.99, 17113, −16070.01, −61921.81, −34733.05</td>
<td>17964.42, 17087, −16209.58, −61905.73, −34758.28</td>
</tr>
<tr>
<td>LCA – 2 class</td>
<td>17946.92, 17087, −16209.58, −61905.73, −34758.28</td>
<td>17762.09, 17051, −16201.57, −61908.02, −34757.39</td>
</tr>
<tr>
<td>FMM – 1 factor 2 class</td>
<td>17976.43, 17089, −16201.57, −61908.02, −34757.39</td>
<td>17762.09, 17051, −16201.57, −61908.02, −34757.39</td>
</tr>
<tr>
<td>FMM – 1 factor 3 class</td>
<td>17762.09, 17051, −16201.57, −61908.02, −34757.39</td>
<td>17762.09, 17051, −16201.57, −61908.02, −34757.39</td>
</tr>
</tbody>
</table>

−2LL = −2 × log likelihood.

Table 4
Varimax rotated factor loadings for the best fitting and parsimonious two factor solution for males and females.

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stressed 1 item</td>
<td>0.97, −0.01, 0.87, 0.08</td>
<td>0.79, 0.40, 0.89, 0.24</td>
</tr>
<tr>
<td>2. Used cannabis in dangerous situations</td>
<td>0.66, 0.31, 0.64, 0.22</td>
<td>0.72, 0.30, 0.69, 0.42</td>
</tr>
<tr>
<td>3. Use caused legal problems or traffic accidents</td>
<td>0.68, 0.21, 0.50, 0.11</td>
<td>0.35, 0.51, 0.43, 0.39</td>
</tr>
<tr>
<td>4. Use caused problems with family, friends, or people at work</td>
<td>0.72, 0.30, 0.69, 0.42</td>
<td>0.72, 0.30, 0.69, 0.42</td>
</tr>
<tr>
<td>5. Use caused physical problems, made depressed or very nervous</td>
<td>0.51, 0.35, 0.22, 0.24</td>
<td>0.72, 0.30, 0.69, 0.42</td>
</tr>
<tr>
<td>6. Failure to fulfill major role obligation at work school or homea</td>
<td>0.84, 0.21, 0.86, 0.24</td>
<td>0.84, 0.21, 0.86, 0.24</td>
</tr>
<tr>
<td>7. When started use ended up taking more than planned</td>
<td>0.52, 0.59, 0.62, 0.56</td>
<td>0.52, 0.59, 0.62, 0.56</td>
</tr>
<tr>
<td>8. Tried to cut down or stop using it</td>
<td>0.00, 0.32, 0.07, 0.76</td>
<td>0.00, 0.32, 0.07, 0.76</td>
</tr>
<tr>
<td>9. Spent time taking or using it, recovering from it, or doing whatever to get it</td>
<td>0.79, 0.40, 0.89, 0.24</td>
<td>0.79, 0.40, 0.89, 0.24</td>
</tr>
<tr>
<td>10. Used so often instead of work, hobbies or spending time with family or friends</td>
<td>0.77, 0.45, 0.88, 0.30</td>
<td>0.77, 0.45, 0.88, 0.30</td>
</tr>
<tr>
<td>11. Needed to use a lot more to get high or feel effects than when first used</td>
<td>0.66, 0.45, 0.67, 0.36</td>
<td>0.66, 0.45, 0.67, 0.36</td>
</tr>
<tr>
<td>12. Withdrawal symptomsb</td>
<td>0.60, 0.52, 0.72, 0.35</td>
<td>0.60, 0.52, 0.72, 0.35</td>
</tr>
</tbody>
</table>

η1 and η2 denote latent factors 1 and 2. Factor loadings >0.40 are bolder for ease of interpretation.

a Aggregation of used often when doing something important OR stayed away from school or missed appointments because of use.

b Aggregation of felt sick when cutting down or stopped use OR after not using cannabis, used to prevent sickness.

Contrary to the uncertainty surrounding the clinical significance of withdrawal symptoms (American Psychiatric Association, 1994), there is converging evidence from laboratory and clinical studies (Budney et al., 2004, 2003; Lichtman and Martin, 2002; Lichtman et al., 2005) that withdrawal diagnostic criteria from part of a valid syndrome. Our pattern of factor loadings strongly capitalized on the twin data to model the genetic etiology of this CUD factor. Finally, although not part of the central focus of this paper we capitalized on the twin data to model the genetic etiology of this CUD factor. And consistent with previous findings showing that familial aggregation in cannabis use, abuse and dependence is best explained by a varying combinations (depending on the measure) of additive genetic and shared environmental variance (Agrawal and Lynskey, 2006; Gillespie et al., 2009) our results shown in Table 5 demonstrate that over half the total variance in the first CUD factor was explained by additive genetic effects. Based on the 95% confidence intervals we found marginal evidence of shared environmental variance with the remainder of the variance accounted

Table 5
Univariate additive genetic (A), shared environmental (C) and non-shared environmental (E) components of variance for latent factor 1 (η1) and latent factor 2 (η2) in the best fitting LFA solution for males and females.

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>η1</td>
<td>0.54 (0.33–0.77)</td>
<td>0.25 (0.03–0.44)</td>
</tr>
<tr>
<td>η2</td>
<td>0.20 (0.00–0.47)</td>
<td>0.01 (0.00–0.10)</td>
</tr>
<tr>
<td>A (95%CI)</td>
<td>0.54 (0.33–0.77)</td>
<td>0.25 (0.03–0.44)</td>
</tr>
<tr>
<td>C (95%CI)</td>
<td>0.20 (0.00–0.47)</td>
<td>0.01 (0.00–0.10)</td>
</tr>
<tr>
<td>E (95%CI)</td>
<td>0.30 (0.00–0.60)</td>
<td>0.16 (0.00–0.55)</td>
</tr>
</tbody>
</table>

Estimates of A, C and E include 95% confidence intervals for each factor are based on univariate models fitted to the bolded items in Table 4 with factor loadings ≥0.40.
for by non-shared or unique environmental variance. Similarly large genetic variance has been reported elsewhere (Kendler and Prescott, 1998; Lynskey et al., 2002; Rhee et al., 2003; van den Bree et al., 1998). The less interpretable second factor for females and the ‘dependence’ factor for males was largely explained by non-shared environmental.

4.1. Limitations

Our findings must be interpreted in the context of at least four potential limitations. First, our data were restricted to a white Virginian sample. Previous analyses based on the Mid Atlantic Twin Registry (Kendler et al., 2000) have suggested that it does not differ from the general population in rates of psychopathologic conditions, including illicit substance-use, and although the prevalence rates cannot be generalized to ethnic minorities and African Americans, it is likely to be broadly representative of white North American men and women. Two, assessments were based on a single interview which necessarily included measurement error. Three, endorsement rates for one item “loss of control, unable to stop…tried to cut down” was extremely low which may have resulted in unstable parameter estimates. Finally, because the current analyses assumed that members of twin pairs were independent, there is the limitation of non-independent data. Failure to take account of statistical non-independence of the twins is not expected to change parameter estimates but may alter their confidence intervals. However, 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.

5. Conclusion

Results from exhaustive model fitting revealed that the abuse and dependence symptoms as well as the stem loaded onto a single, general factor. Although a second but less interpretable factor that partly captured the dependence and withdrawal symptoms was observed in males, any incremental contribution made by this second factor was likely to be minimal. These conclusions are commensurate with an emerging consensus for a single factor to explain the covariance between use and diagnostic criteria. In addition, these support the proposed changes in the forthcoming DSM-V with regard to substance-related disorders.1

Before attempting to replicate these findings using other nationally representative samples, our next aim will be to investigate sex differences in the observed factors within a measurement invariance framework (Meredith, 1993; Vandenberg, 2000). This will enable quantification and significance testing of possible measurement differences in the factor means, factor loadings or item thresholds, as well as correction for these effects if found. Following this, we will further leverage this genetically informative data to estimate maximally heritable factor scores for each individual in order to provide increased power to detect quantitative trait loci signals for CUD and correlated drug use phenotypes as part of anticipated genome wide tests of genomic association.

Role of funding source

Funding was received from the US National Institute on Drug Abuse (R00DA023549, 1K99DA023549-01A2, DA-18673) and the US National Institutes of Health (DA-11287, MH/AA/DA-49492, MH-01458, AA-00236). The US NIH and NIDA had no role in the design of the study, data collection, data analysis and interpretation of data, in the writing of this manuscript, or decision to submit this manuscript for publication.

Contributors

Dr. Gillespie designed the study, undertook the statistical analyses, reviewed the literature, and wrote the first draft manuscript. Dr. Neale assisted in the design and implementation of the statistical analyses. All authors contributed to and approved the final manuscript.

Conflict of interest

There are no conflicts of interest.

Acknowledgements

The authors thank Ms. Indrani Ray for database assistance and Dr. Linda Corey for assistance with the ascertainment of twins from the Virginia Twin Registry, now part of the Mid-Atlantic Twin Registry (MATR), directed by Dr. Judy Silberg. The MATR has received support from NIH, the Carman Trust, and the W.M. Keck, John Templeton, and Robert Wood Johnson Foundations.

References


Murphy, B., 2006. Should substance use disorders be considered as categorical or dimensional? Addiction 101, 616–117.


