Factor and item-response analysis DSM-IV criteria for abuse of and dependence on cannabis, cocaine, hallucinogens, sedatives, stimulants and opioids

Nathan A. Gillespie1, Michael C. Neale1, Carol A. Prescott1, Steven H. Aggen1 & Kenneth S. Kendler1,2

Department of Psychiatry1 and Department of Human Genetics, Virginia Institute of Psychiatric and Behavioral Genetics, Virginia Commonwealth University, VA, USA2

ABSTRACT

Aims This paper explored, in a population-based sample of males, the factorial structure of criteria for substance abuse and dependence, and compared qualitatively the performance of these criteria across drug categories using item–response theory (IRT). Design Marginal maximum likelihood was used to explore the factor structure of criteria within drug classes, and a two-parameter IRT model was used to determine how the difficulty and discrimination of individual criteria differ across drug classes. Participants A total of 4234 males born from 1940 to 1974 from the population-based Virginia Twin Registry were approached to participate. Measurements DSM-IV drug use, abuse and dependence criteria for cannabis, sedatives, stimulants, cocaine and opiates. Findings For each drug class, the pattern of endorsement of individual criteria for abuse and dependence, conditioned on initiation and use, could be best explained by a single factor. There were large differences in individual item performance across substances in terms of item difficulty and discrimination. Cocaine users were more likely to have encountered legal, social, physical and psychological consequences. Conclusions The DSM-IV abuse and dependence criteria, within each drug class, are not distinct but best described in terms of a single underlying continuum of risk. Because individual criteria performed very differently across substances in IRT analyses, the assumption that these items are measuring equivalent levels of severity or liability with the same discrimination across different substances is unsustainable. Compared to other drugs, cocaine usage is associated with more detrimental effects and negative consequences, whereas the effects of cannabis and hallucinogens appear to be less harmful. Implications for other drug classes are discussed.

Keywords Cannabis, cocaine, dependence, DSM-IV, hallucinogens, item–response theory, sedatives, stimulants, opioids, substance abuse.

Correspondence to: Nathan A Gillespie, Virginia Institute of Psychiatric and Behavioral Genetics, Department of Psychiatry and Department of Human Genetics, Virginia Commonwealth University, 800 East Leigh Street, Biotech 1, Suite 101, Richmond, VA 23219–1534, USA. E-mail: ngillespie@vcu.edu Submitted 2 May 2006; initial review completed 14 September 2006; final version accepted 18 December 2006

INTRODUCTION

DSM-III and its successors have made two important assumptions in the assessment of psychoactive substance use disorders: (i) individual criteria for abuse and dependence define two distinct dimensions of drug-related problems; and (ii) the same criteria for abuse and dependence are applicable to all classes of psychoactive drugs.

One empirical way to assess the relationship between criteria assessing abuse and dependence is to submit them to factor-analysis. Factor-analytical approaches have typically followed two forms of enquiry: are drug dependence criteria unidimensional; and do the combined dependence and abuse criteria yield the two predicted dimensions across drug categories? Five studies [1–5] have followed this first line of enquiry by examining DSM-III-R [6] and DSM-IV [7] dependence criteria only. Despite variation in sample size, the alcohol, cocaine and opiate dependence criteria demonstrated consistently a unidimensional structure, whereas the cannabis dependence criteria demonstrated a unidimensional structure in only three of the four studies. Hallucinogen criteria did...
not demonstrate a unidimensional structure, and results for stimulants and sedatives remain equivocal. Five other studies [8–12] have followed the second approach, whereby the abuse and dependence criteria were analysed jointly. Based on a mixture of clinical and population-based samples, which were administered DSM-IV criteria, this method has found consistently that dependence and abuse criteria load onto a single factor. These results are surprising. Unlike the DSM-III-R, which included overlapping criteria to define abuse and dependence, the DSM-IV was redefined intentionally to ensure non-overlapping criterion sets for abuse and dependence. It is also worth noting that a cannabis withdrawal factor was not extracted in any of these studies. Despite accumulating evidence showing that cannabis withdrawal is a valid syndrome [13], the DSM-IV excluded withdrawal from the diagnosis of cannabis dependence.

In addition to questions of dimensionality, another important empirical question is whether individual criteria perform similarly across different drug classes. Their relative performance is of interest because drugs are likely to differ in their physiological, behavioural and cognitive effects. They are also likely to differ in their legal status, pharmacological effects and addictive potential [2], as well as their pattern of absorption and metabolism and their rewarding and dysphoric properties. Given these differences, it seems unlikely that the response characteristics for each item will remain invariant across drug classes. Item–response theory (IRT) is an appropriate method for answering these sorts of questions.

Briefly, IRT can be used to evaluate the psychometric properties of items or assessment criteria [14]. The primary item characteristics provided by this method are (i) where on the liability scale an item has a 50:50 chance of endorsement (i.e. its difficulty), and (ii) how rapidly this chance of endorsement changes across the liability dimension (i.e. the item’s discrimination). These characteristics make IRT relevant for assessing abuse and dependence criteria not just within drug classes, but also across substances. Studies of these properties have the potential to lead to better scale construction, through the elimination of items that are redundant or that fail to discriminate adequately. As discussed by Takane & DeLeeuw [15], there is a direct equivalence of the normal ogive item–response model to the factor-analysis model for binary data. Essentially, the item difficulty parameters correspond to the item thresholds, while the discrimination parameters correspond to the factor loadings. Previously, IRT approaches have been used to study the construct validity of DSM-IV abuse and dependence criteria for alcohol, cannabis and cocaine [8,16]. Langenbucher and colleagues found that the abuse and dependence criteria for alcohol, cannabis and cocaine did not function as DSM-IV intended; not only did no clear distinction between the two syndromes emerge, but the IRT results suggested that the criteria were undersampling extreme levels of pathology while oversampling moderate ones.

This paper has two aims. The first is to explore, in a large population-based sample of males, the factor structure of DSM-IV criteria for psychoactive substance abuse and dependence for six major classes of illicit psychoactive drugs (cannabis, cocaine, stimulants, sedatives, opiates and hallucinogens). The second is to compare the item–response characteristics of abuse and dependence criteria across drug categories using an IRT approach. Unlike previous studies, our approach will be novel in that our IRT parameter estimates will be conditioned on drug initiation and use.

**METHODS**

Sample and assessment procedures

This report is based on data collected in a study of Caucasian adult twins from the Virginia Twin Registry. This registry was formed by a systematic search of all Virginia birth certificates since 1918. Further details are available elsewhere [17,18]. The interview which collected the current data was completed by 5642 (82.6%; males = 4234, females = 1408) of those who had cooperated with a first interview 1 year previously. Where possible, this interview was completed face to face (for 79.4% of sample). Signed informed consent was obtained before all face-to-face interviews and verbal consent before all telephone interviews.

Analyses were based on the male data only. Subjects were an average age of 34 years (SD = 9.26, range = 19–57 years) and had a mean of 13.6 years of education (SD = 2.6). Interviewers had a master’s degree in a mental health-related field or a bachelor’s degree in this area plus 2 years of clinical experience.

The interview included assessments of life-time drug use, abuse and dependence across six categories of substances using an adaptation of the Structured Clinical Interview (SCID) for DSM-III-R [19]. The categories and common examples were cannabis (marijuana and hashish); sedatives (quaalude, Seconal and Valium); stimulants (speed, ecstasy and Ritalin); cocaine (intranasal, freebase and crack); opiates [heroin, Demerol (meperidine hydrochloride) and morphine]; and hallucinogens (lysergic acid diethylamide, mescaline and phenycyclidine). For substances that could be obtained legally, we defined non-medical use as use (1) without a doctor’s prescription, (2) in greater amounts or more often than prescribed or (3) for any other reason than a doctor said it should be taken. For each drug class, the abuse and dependence diagnoses comprised 13 items listed in Table 1. This included (1) tolerance and (2) withdrawal...
Table 1  Marginal maximum likelihood (MML) endorsements\(^1\) (%) for the stem (use) and probe (abuse and dependence) items based on subjects who reported illicit drug use, experimentation or non-medical use of licit substances.

<table>
<thead>
<tr>
<th>Have you ever taken any of these drugs?(^2)</th>
<th>Cannabis</th>
<th>Cocaine</th>
<th>Stimulants</th>
<th>Sedatives</th>
<th>Opiates</th>
<th>Hallucinogens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endorsement of stem items (administration conditional on taking drug)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 = neither 11 or more times in a month or 6 or more times in your life</td>
<td>36.2 (29.7)</td>
<td>40.7 (37.0)</td>
<td>41.7 (33.6)</td>
<td>53.3 (44.3)</td>
<td>44.2 (40.8)</td>
<td>57.3 (48.7)</td>
</tr>
<tr>
<td>1 = 6 or more times in your life</td>
<td>44.5 (39.7)</td>
<td>44.1 (43.9)</td>
<td>42.0 (44.4)</td>
<td>32.7 (43.0)</td>
<td>34.8 (35.7)</td>
<td>33.7 (44.2)</td>
</tr>
<tr>
<td>2 = 11 or more times in a month and (by default) 6 or more times in your life</td>
<td>19.3 (30.6)</td>
<td>15.1 (19.1)</td>
<td>16.3 (22.0)</td>
<td>14.0 (12.7)</td>
<td>21.0 (23.5)</td>
<td>9.0 (7.2)</td>
</tr>
<tr>
<td>Endorsement of probe items (administration conditional on use and stem score = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hazardous use</td>
<td>28.9 (37.4)</td>
<td>39.4 (44.0)</td>
<td>32.5 (40.1)</td>
<td>30.1 (42.1)</td>
<td>31.8 (41.9)</td>
<td>30.4 (44.6)</td>
</tr>
<tr>
<td>Consequences: legal</td>
<td>2.8 (3.9)</td>
<td>5.3 (7.0)</td>
<td>1.7 (2.9)</td>
<td>3.1 (5.7)</td>
<td>3.6 (6.4)</td>
<td>0.9 (1.8)</td>
</tr>
<tr>
<td>Consequences: social</td>
<td>9.7 (13.8)</td>
<td>19.6 (24.1)</td>
<td>6.3 (9.1)</td>
<td>8.2 (14.6)</td>
<td>10.4 (17.0)</td>
<td>6.1 (11.3)</td>
</tr>
<tr>
<td>Consequences: physical &amp; psychological</td>
<td>9.3 (12.3)</td>
<td>26.1 (30.4)</td>
<td>23.1 (29.3)</td>
<td>7.4 (12.4)</td>
<td>10.5 (17.4)</td>
<td>13.3 (21.2)</td>
</tr>
<tr>
<td>Used often when doing something . . . important</td>
<td>15.7 (22.1)</td>
<td>13.3 (16.6)</td>
<td>33.3 (40.6)</td>
<td>12.0 (17.8)</td>
<td>17.1 (24.6)</td>
<td>5.9 (11.7)</td>
</tr>
<tr>
<td>Stay away from school/miss appointments</td>
<td>6.0 (8.6)</td>
<td>13.2 (16.8)</td>
<td>3.7 (5.7)</td>
<td>7.2 (12.4)</td>
<td>8.3 (12.8)</td>
<td>6.5 (11.6)</td>
</tr>
<tr>
<td>Used more or longer than thought/planned</td>
<td>11.1 (16.1)</td>
<td>32.6 (39.9)</td>
<td>5.1 (22.2)</td>
<td>10.2 (18.0)</td>
<td>15.7 (24.3)</td>
<td>6.1 (12.4)</td>
</tr>
<tr>
<td>Loss of control: unable to stop/desire to stop, tried to cut down or stop using it</td>
<td>3.7 (4.4)</td>
<td>14.3 (14.3)</td>
<td>8.1 (5.5)</td>
<td>2.2 (5.3)</td>
<td>3.8 (8.1)</td>
<td>2.9 (4.7)</td>
</tr>
<tr>
<td>Spend time taking/using it, recovering from it, or doing whatever</td>
<td>9.0 (13.7)</td>
<td>21.9 (26.7)</td>
<td>2.7 (12.2)</td>
<td>7.4 (13.4)</td>
<td>13.8 (21.9)</td>
<td>6.9 (13.9)</td>
</tr>
<tr>
<td>Used instead of work/hobbies</td>
<td>6.1 (7.9)</td>
<td>15.2 (19.5)</td>
<td>13.0 (4.3)</td>
<td>4.1 (8.5)</td>
<td>7.0 (14.8)</td>
<td>3.8 (8.4)</td>
</tr>
<tr>
<td>Need for larger amounts/doses (tolerance)</td>
<td>13.1 (19.1)</td>
<td>25.4 (30.5)</td>
<td>8.6 (18.5)</td>
<td>7.6 (14.3)</td>
<td>11.5 (19.5)</td>
<td>5.5 (11.8)</td>
</tr>
<tr>
<td>Withdrawal symptoms: feeling sick when cutting down/stopping</td>
<td>2.5 (3.5)</td>
<td>17.0 (21.3)</td>
<td>3.3 (12.7)</td>
<td>4.8 (9.5)</td>
<td>8.8 (16.0)</td>
<td>1.6 (4.3)</td>
</tr>
<tr>
<td>Withdrawal symptoms: after not using . . . use to prevent sickness</td>
<td>1.2 (1.7)</td>
<td>6.0 (8.4)</td>
<td>5.1 (5.2)</td>
<td>2.8 (6.1)</td>
<td>6.2 (12.4)</td>
<td>0.3 (0.9)</td>
</tr>
</tbody>
</table>

\(^1\)Item endorsements represent the standard normal cumulative distribution based on MML threshold estimates. The estimated proportion of individuals who endorse an item, conditional on use, when calculated under MML will differ from the sample statistic (italics) because the location of item’s threshold is sensitive to the correlation between the stem and probes, and this correlation will differ across substances. *For substances that could be legally obtained, we defined non-medical use as use (i) without a doctor’s prescription, (ii) in greater amounts or more often than prescribed or (iii) for any other reason than a doctor said it should be taken.

questions from the dependence criteria. The items were structured in this way to permit assignment of both DSM-III-R and DSM-IV diagnostic definitions. Initially, all 13 items were measured on a three-point scale (definite/probable/no). For the analyses reported here ‘probable’ responses were combined with ‘definite’ and the data were analysed dichotomous outcomes.

In order to identify eligible participants, subjects were first asked if they had ever taken medicines or drugs for non-medical purposes. If endorsed, the administration of the drug abuse and dependence module was then contingent upon each subject’s responses to two binary (0/1) stem items. Subjects who endorsed using a drug ‘six or more times in a life-time’ but not ‘11 times in a month’ were asked only the first six abuse items in Table 1. If any abuse items were endorsed, the dependence items were also asked. Subjects who endorsed using a drug ‘11 times in a month’ were automatically asked all 13 items. In order to determine whether these two stem items (‘six or more times in a lifetime’ and ‘11 times in a month’) were an effective screening tool, we combined them for each drug class into a three-point ordinal scale (0/1/2) and included them in the factor-analyses. Table 1 includes the number of subjects with ordinal scale (0/1/2) and included them in the factor-analyses.

Statistical methods

Raw ordinal data analysis

Analysis of our ordinal data assumes that the observed ordinal categories within each item are an imprecise measure of a latent normal distribution of liability, and that this liability distribution has one or more threshold values which discriminate between the categories.
Figure 1  Hypothetical item–response curves (IRCs). The IRC (i) on the far left in Fig. 1 is the least difficult, but has approximately the same level of discrimination as the second curve (ii), as both have equivalent slopes. The third IRC (iii) is the flattest and is consequently the least discriminating because the probability of endorsing a symptom \( P(\theta) \) changes the slowest as the liability \( \theta \) increases. As an example of perfect discrimination, the fourth IRC curve (iv), discriminates perfectly between subjects who fall above and below a liability or risk of \( \theta = 2 \), to the left of the vertical line the probability of a correct response is 0 compared to 1 on the right side of the line. When plotting the IRCs, we used a two-parameter logistic model written as \( P(\theta) = 1/1 + e^{(a_\theta - b_\theta)/\lambda} \) to trace the probability of endorsing a symptom \( P(\theta) \) for a person at each level of the latent trait level, where \( b \) and \( a \) denote the difficulty and discrimination parameters, respectively. Although the theoretical range of latent liability \( \theta \) is \( -\infty \) to \( +\infty \), ours extended from \(-3.5\) to \(+3.5\). A constant, 1.7, was also added to the model as an adjustment so that the logistic model approximates the normal ogive thresholds can be conceived of as ‘cut-points’ along a standard normal distribution which classify individuals in terms of a probability or risk of endorsing one of two or more discrete (ordinal) categories. By considering one category at a time, thresholds can be estimated so that the proportion of the distribution lying between adjacent thresholds matches exactly the observed proportion of the sample that is found in each category.

Marginal maximum likelihood (MML) factor-analysis

In order to determine the factor structure, we fitted one-, two- and three-factor solutions (unrotated and Varimax) to the raw data using an MML [20] approach in Mx [21]. The item factor loadings, residuals and thresholds can then be used to calculate a two-parameter normal ogive item–response model.

Item–response modeling

The MML estimate for the thresholds and the factor loadings can then be transformed into the IRT parameters. Figure 1 describes the relationship between an individual’s latent liability or risk \( \theta \) and the probability of endorsing a particular item \( P(\theta) \). The ‘s–shaped’ item–response curves (IRC) are standardized distributional curves which represent the cumulative frequencies for the probability of endorsing an item at increasing levels of liability, which can be described in terms of two parameters—difficulty (b) and discrimination (a) [8]:

\[
\begin{align*}
b_\theta &= \tau_{\theta} - \nu_{\theta}/\lambda_{\theta} \\
a_\theta &= \lambda_{\theta}/\sqrt{\theta}
\end{align*}
\]

Difficulty \( (b_{\theta}) \) refers to the point on the underlying latent liability \( \theta \) at which the probability \( P(\theta) \) of meeting that criterion equals 50% [8]. Discrimination \( (a_{\theta}) \), measures the degree of precision with which items can distinguish between participants at different levels of latent liability ‘above’ versus ‘below’ an item’s threshold [8]. When an item discriminates well, as reflected by a steeper slope, a greater proportion of subjects at high risk are more likely to endorse a particular item compared to subjects at lower levels of risk.

Note that in the context of our item analysis, to obtain estimates of population-level parameters is somewhat different from the IRT scale-scoring context. In the latter, an estimate of an individual’s ability may not be obtained from certain algorithms due to the floor or ceiling effects, such as when all item–responses are identical. However, in the context of characterizing the population, retaining such individuals in the analysis is essential to obtain unbiased population parameters.

Ascertainment bias and cross-drug comparisons

If the latent liability distribution \( \theta \) differs across classes, an important question is whether we can compare item difficulty and discrimination validly across drug classes. An advantage of ML estimates is that they are robust to certain forms of missing data. In the present case, the ordinal stem item for initiation predicts whether or not the probe items for abuse and dependence symptoms are missing. Therefore, despite the relatively large proportion of ‘missingness’ for these items, because we have included the ordinal stem in our analyses MML can, in effect, correct for the fact that we are missing data on abuse and dependence from subjects who denied all the stem items. It is therefore reasonable to expect good recovery of the population values of the parameters.

By including the ordinal stem, we can therefore pose the question ‘how well does criterion x measure the latent trait or liability to develop symptoms of abuse and dependence for a particular substance?’. However, a problem arises because the liability to initiate use and the liability to develop a particular symptom of abuse or dependence may correlate, and this correlation may differ between substances.

Fortunately, ML estimates produce asymptotically unbiased estimates of (i) the proportion of people in the population who would develop the symptom, if they were
to initiate use, i.e. the threshold; and (ii) of the correlation between liability to initiate (LI) and the liability to a dependence symptom (LDS), represented as the factor loading. Hence the importance of including the ordinal stem item in the analyses. Comparisons across substances in terms of item difficulty and discrimination become feasible, despite the fact that their LI and LDS correlations may differ (see Neale and colleagues [22] for a more complete treatment of this statistical issue).

RESULTS

Maximum likelihood estimates of the proportion of people in the population who endorse the abuse and dependence criteria, conditional on use more than six times in a life-time, are shown in Table 1. Also shown for comparison are the (asymptotically biased) standard sample item endorsements for the abuse and dependence criteria conditional on use but not adjusted for any correlation between stem and probe items.

MMI factor-analysis

The eigenvalues for each substance, which were calculated on a pairwise basis from the matrix of polychoric correlations, indicated a two-factor solution for several substances. The first three eigenvalues for cannabis were 7.92, 1.20 and 0.94. Similar patterns of high first to second eigenvalue ratio were observed for cocaine (10.46, 0.76, 0.64), stimulants (8.49, 1.07, 0.86), sedatives (10.11, 0.90, 0.72), opioids (10.61, 0.99, 0.79) and hallucinogens (9.30, 1.98, 0.85). For cannabis and stimulants, both withdrawal symptoms ‘feeling sick when cutting down/stopping’ and ‘after not using . . . use to prevent illness’ loaded highly onto the second factor in a two-factor solution. For hallucinogens, only one withdrawal symptom, ‘feeling sick when cutting down/stopping’, loaded highly onto the second factor in a two-factor solution.

The Akaike information criterion (AIC) and Bayesian information criterion (BIC) suggested that the two- and three-factor solutions provided a better fit to the data compared to a unidimensional model. While the fit indices and eigenvalues indicated either a two- or three-factor solution, these solutions were far less interpretable than a one-factor model. Items with substantial loadings on the second and third factors, in all instances, still had larger factor loadings on the first, general factor and neither Varimax nor Promax rotations improved the interpretability of the two- and three-factor models.

It is important to note that more saturated models will tend to provide better fits when based on large sample sizes such as the current one, and so a one-factor solution was chosen for the all drug classes. As an added check, we fitted a model in which the abuse and dependence criteria were allowed to load only onto the first and second factors, respectively. Although the more saturated two-factor model provided a slightly better fit to the data in terms of AIC and BIC for all drug classes when compared to a single-factor solution, the latent factor correlations were all extremely high: cannabis \( r = 0.98 \); cocaine \( r = 0.95 \); stimulants \( r = 0.92 \); sedatives \( r = 0.95 \); opioids \( r = 0.93 \); and hallucinogens \( r = 0.98 \). This suggested that the overwhelming majority of variance in both abuse and dependence symptoms is accounted for by a single liability dimension for all drug classes. Maximum likelihood parameter estimates from this single-factor model, along with their 95% bootstrapping confidence intervals, are shown in Table 2.

Item–response curves

Given evidence for a single dimension of risk underlying all the abuse and dependence criteria, we proceeded to fit IRCs for these criteria to each drug class, as shown in Fig. 2.

The first criterion, ‘hazardous use’, appears to be assessing approximately the same levels of difficulty with the same degree of discrimination across all six drug classes.

The ‘legal consequences’ item discriminates less well among hallucinogen users, but better for stimulant users. In terms of difficulty, large differences are seen for the difficulty of this item between drug classes. Higher risk is required before legal consequences emerge hallucinogens. In contrast, lower risk is required before legal consequences are observed among cocaine users. The remaining drugs are intermediate.

‘Social consequences’ are more likely to be endorsed by cocaine users at lower risk than for other substances. On the other hand, higher risk is required before stimulant and hallucinogen users manifest these problems. No major differences are seen in terms of discrimination.

‘Physical and psychological symptoms’ are also endorsed by cocaine and stimulant users at lower risk than other drug classes, whereas the same symptoms manifest among cannabis users at higher risk. The flatness of slope also reveals very poor item discrimination for cannabis, stimulant and hallucinogen users.

The most striking feature with ‘used often when doing something . . . important’ is the difference in the curve for stimulant users versus all other substances. Much lower liability is required for stimulant users to endorse ‘doing something important’ than with the other drug users. By contrast, sedative and hallucinogen users tend to endorse this item only at higher risk.
Staying away from school/miss appointments manifests at similar levels of risk across the drugs except for cocaine; at lower liability this group is more likely to endorse this item.

In terms of individuals who fall either above or below the threshold for symptom endorsement ‘used more or longer than thought/planned’ discriminates well among cocaine users. However, item discrimination must be interpreted with caution because of the low endorsement for this criterion. As for difficulty, hallucinogen and cannabis users, followed by sedative users, are unlikely to endorse this symptom except at higher levels of risk. This is in contrast to cocaine, whose users are more likely to endorse this item at lower risk versus hallucinogen users at higher risk. Discrimination and difficulty parameters for stimulant and cannabis users are almost identical.

Table 2 Estimated factor loadings (and 95% bootstrapping confidence intervals) for the abuse and dependence items for each drug category from the adapted structured clinical interview.

<table>
<thead>
<tr>
<th>Item</th>
<th>Cannabis F1</th>
<th>Cocaine F1</th>
<th>Stimulants F1</th>
<th>Sedatives F1</th>
<th>Opiates F1</th>
<th>Hallucinogens F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hazardous use</td>
<td>0.69 (0.64–0.74)</td>
<td>0.75 (0.69–0.82)</td>
<td>0.68 (0.61–0.76)</td>
<td>0.77 (0.66–0.90)</td>
<td>0.65 (0.46–0.81)</td>
<td>0.74 (0.62–0.85)</td>
</tr>
<tr>
<td>Consequences: legal</td>
<td>0.64 (0.54–0.72)</td>
<td>0.71 (0.60–0.81)</td>
<td>0.82 (0.72–0.94)</td>
<td>0.68 (0.48–0.87)</td>
<td>0.62 (0.39–0.85)</td>
<td>0.56 (0.02–1.0)</td>
</tr>
<tr>
<td>Consequences: social</td>
<td>0.73 (0.68–0.78)</td>
<td>0.83 (0.78–0.88)</td>
<td>0.71 (0.58–0.83)</td>
<td>0.82 (0.70–0.95)</td>
<td>0.77 (0.64–0.89)</td>
<td>0.69 (0.55–0.84)</td>
</tr>
<tr>
<td>Consequences: physical and psychological</td>
<td>0.54 (0.46–0.63)</td>
<td>0.77 (0.70–0.83)</td>
<td>0.63 (0.54–0.74)</td>
<td>0.72 (0.57–0.87)</td>
<td>0.85 (0.71–0.96)</td>
<td>0.64 (0.52–0.77)</td>
</tr>
<tr>
<td>Used often when doing something . . . important</td>
<td>0.77 (0.73–0.81)</td>
<td>0.74 (0.63–0.82)</td>
<td>0.61 (0.54–0.71)</td>
<td>0.60 (0.47–0.72)</td>
<td>0.64 (0.51–0.82)</td>
<td>0.78 (0.69–0.87)</td>
</tr>
<tr>
<td>Stay away from school/miss appointments</td>
<td>0.75 (0.69–0.80)</td>
<td>0.79 (0.73–0.85)</td>
<td>0.76 (0.61–0.87)</td>
<td>0.74 (0.63–0.85)</td>
<td>0.59 (0.44–0.78)</td>
<td>0.66 (0.55–0.80)</td>
</tr>
<tr>
<td>Used more or longer than thought/planned</td>
<td>0.80 (0.77–0.84)</td>
<td>0.99 (0.87–1.0)</td>
<td>0.84 (0.79–0.90)</td>
<td>0.89 (0.83–0.99)</td>
<td>0.92 (0.83–1.0)</td>
<td>0.98 (0.88–1.0)</td>
</tr>
<tr>
<td>Loss of control: unable to stop/desire to stop, tried to cut down or stop using it</td>
<td>0.75 (0.69–0.81)</td>
<td>0.88 (0.79–0.96)</td>
<td>0.73 (0.62–0.87)</td>
<td>0.93 (0.83–1.0)</td>
<td>0.88 (0.77–1.0)</td>
<td>0.80 (0.64–0.95)</td>
</tr>
<tr>
<td>Spend time taking/using it, recovering from it, or doing whatever</td>
<td>0.89 (0.85–0.93)</td>
<td>0.88 (0.84–0.93)</td>
<td>0.88 (0.80–0.97)</td>
<td>0.87 (0.78–0.99)</td>
<td>0.87 (0.79–0.99)</td>
<td>0.84 (0.76–0.93)</td>
</tr>
<tr>
<td>Used instead of work/hobbies</td>
<td>0.99 (0.93–1.0)</td>
<td>0.88 (0.83–0.93)</td>
<td>0.83 (0.67–0.98)</td>
<td>0.89 (0.78–1.0)</td>
<td>0.99 (0.90–1.0)</td>
<td>0.85 (0.78–0.96)</td>
</tr>
<tr>
<td>Need for larger amounts/doses (tolerance)</td>
<td>0.83 (0.79–0.87)</td>
<td>0.86 (0.82–0.94)</td>
<td>0.83 (0.78–0.89)</td>
<td>0.90 (0.82–1.0)</td>
<td>0.93 (0.88–1.0)</td>
<td>0.87 (0.78–0.99)</td>
</tr>
<tr>
<td>Withdrawal symptoms: feeling sick when cutting down/stoping</td>
<td>0.78 (0.71–0.85)</td>
<td>0.85 (0.80–0.90)</td>
<td>0.81 (0.75–0.89)</td>
<td>0.86 (0.76–0.99)</td>
<td>0.87 (0.80–1.0)</td>
<td>0.93 (0.84–1.0)</td>
</tr>
<tr>
<td>Withdrawal symptoms: after not using . . . use to prevent sickness</td>
<td>0.74 (0.66–0.82)</td>
<td>0.79 (0.70–0.89)</td>
<td>0.82 (0.73–0.92)</td>
<td>0.93 (0.85–1.0)</td>
<td>0.98 (0.96–1.0)</td>
<td>0.87 (0.36–1.0)</td>
</tr>
</tbody>
</table>

F1 = factor loadings.
DISCUSSION

Factor structure

Our first goal was to determine the factor structure of DSM-IV criteria for the diagnosis of psychoactive substance abuse and dependence. Eigenvalues for cannabis and cocaine were in the range of those reported elsewhere [8]. Edward’s [23] bi-axial concept of the independence of abuse and dependence was not supported by our data. Instead, we found that the DSM-IV abuse and dependence items were best explained by a unidimensional construct.

Figure 2  Item–response curves for the DSM-IV abuse and dependence criteria > Note: The S-shaped item response curves are obtained by using the cumulative normal distribution function to estimate the probability of endorsing an item at increasing levels of liability. Each curve describes the relationship between an individual’s liability ($\theta$) on the x-axis, and the probability of endorsing an item or criterion [$P(\theta)$] on the y-axis. Although not shown, the theoretical range in liability extends from $-\infty$ to $+\infty$. In a two parameter model such as this, all item response curves asymptote to probabilities of 0 and 1 on the left and right tails respectively.
In an analysis of DSM-IV dependence and abuse criteria, Nelson and colleagues [10] also reported a unidimensional structure. However, they argue that such results can be biased by the large numbers of non-clinical and severe clinical respondents within population-based samples. When they limited their analyses to respondents with low to moderate item endorsement, a two-factor solution ‘dependence’ and ‘abuse’ provided a better fit for cannabis and opioids. We did not restrict our sample. Our findings are consistent with those of Feingold & Rounsaville [11,12], who found that a one-factor model provided a better fit for alcohol, cannabis, cocaine and sedatives (but not opiates). Feingold & Rounsaville [11,12] also reported that the DSM-IV abuse and dependence criterion correlated with the Addiction Severity Index [24] score, which assesses consequences of substance abuse. This is inconsistent with the prediction that abuse and dependence criteria are qualitatively distinct. Our results suggest that the physiological, behavioural and cognitive problems related to psychoactive substance cluster into a single dimension. This is the same for all drug types and reflects a continuum of risk. If dependence represents a more severe phenotype than abuse, as has been suggested [11], we would expect to see lower estimates of item difficulty for the abuse items. We found no such trend for any substance. Our findings argue against maintaining the current DSM-IV distinction between abuse and dependence.

**Item–response curves**

Our IRCs reveal that individual criteria assess differentially the levels of risk and have different powers of discrimination across the drug classes. We review these results by drug class.

**Cannabis**

Compared to other drugs, the effects of cannabis use are amongst the least disruptive. Users are less likely to endorse hazardous use, and are least likely to experience physical and psychological consequences. Along with hallucinogen users, they are also less prone to encounter legal consequences. Compared to other drugs, Kosten et al. [1] have shown that cannabis use is associated with fewer dependence symptoms. Despite previous research [5], which has shown that withdrawal symptoms such as desire or attempts to control use are commonly reported, cannabis users in this sample begin to manifest these DSM-IV criteria only at much higher levels of risk than seen with other substances. The DSM-III-R warns that substances such as cannabis and hallucinogens may not have a clear withdrawal syndrome. This implies that the withdrawal criteria may not be applicable. Smith [25] argues that there is no strong evidence to support a clear cannabis withdrawal syndrome. Moreover, the long half-life of tetrahydrocannabinol (THC) means that all but the heaviest users are unlikely to experience withdrawal, even after abrupt cessation. Despite this, there is now evidence to suggest that THC does have an identifiable withdrawal pattern [13].

**Cocaine**

Our results show that cocaine is the most disabling drug class. Compared to other drug classes, cocaine users are more likely to endorse items at much lower levels of latent liability. Our results are also consistent with the observation that cocaine users are particularly susceptible to ‘crashes’, even following limited recreational use [26]. Cocaine users more likely to have encountered legal, social, physical and psychological consequences, signs of dependence such as time spent using and recovering from use, using more than planned and losing control manifest at lower risk. But as with opioid users, cocaine is also characterized by greater disruption to work and hobbies and a need for increasing amounts, as well as withdrawal symptoms.

**Stimulants**

Although this drug class includes licit varieties which can be used to increase alertness and productivity as well as reduce weight [27–29], it is important to note that our focus was on non-medical use as defined in the Methods. One of the more striking results in this study is the endorsement, at lower risk, of stimulant use when doing something important. Although this item does not discriminate well compared to other drug classes, the results suggest that stimulant use often occurs during work and recreation. Perhaps this reflects use to improve alertness. Along with hallucinogens, users are also less likely to endorse social consequences. However, stimulant and cocaine users are particularly likely to report physical and psychological consequences at comparatively lower levels of risk.

**Sedatives**

The liability among sedatives users to manifest symptoms was often intermediate compared to other drug classes. There were, however, several exceptions. Along with cannabis, sedatives users were unlikely to endorse physical and psychological consequences except at higher risk. Sedative and hallucinogen users were also unlikely to take these drugs while doing something important, and did not require increasingly larger doses, except at higher levels of risk.

**Opioids**

Apart from inducing physical dependence, the speed and the severity of withdrawal associated with opioids, such
as heroin, occurs quickly because of its extremely short half-life [30]. The need for larger doses to achieve the same euphoria was endorsed at levels of risk similar to stimulant and cannabis users, and at even lower levels by cocaine users. Opioid users are likely to report using the drug to prevent sickness. The fact that this withdrawal symptom is more likely to manifest at lower risk compared to most other drug groups suggests that users are motivated to continue use as means of avoiding withdrawal. Kosten and colleagues [1] noted that users report greater use of opioids than planned compared to other drugs. We found that opioid and stimulant users were just as likely (after cocaine users) to manifest this symptom at similar levels of risk.

**Hallucinogens**

Hallucinogens are among the oldest-known group of drugs and are used for their ability to alter human perception and mood. In contrast to cocaine (for whom loss of control is more easily manifested), hallucinogen users are less likely to endorse needing larger amounts or doses. Evidence suggests that hallucinogen use tends to be more recreational rather than in occupational settings [31]. This is supported by our results. Using hallucinogens when doing something important is unlikely to be endorsed and is also associated with fewer indications of disruption except at high levels of risk. Hallucinogens do not appear to show classic patterns of tolerance and withdrawal liability; very few hallucinogen users experience an inability to cut down or control use or a need for larger doses, which are key indicators of loss of control [1,2,32]. Our results suggest that this class of drug’s physical and psychological consequences are intermediate between cocaine and stimulants, on one hand, and cannabis on the other hand. Similar to cannabis, hallucinogen users are unlikely to endorse feeling sick when withdrawal. However, because of the extremely low endorsement, we are cautious about interpreting and comparing the item–response characteristics of using “hallucinogens to prevent sickness” with other drug classes.

**Limitations**

Our findings must be interpreted in the context of at least seven potential limitations, as follows. (1) Our data were restricted to white males born in Virginia. Although it is unclear whether these results will generalize to females, previous analyses from this sample [17] suggest that it does not differ from the general population in rates of psychopathological conditions, including illicit substance use, and that it is likely to be broadly representative of US men. (2) Our analyses did not include cohort effects. Preliminary analyses with age at interview and age of first drug use as covariates for each drug class suggest no significant changes. (3) Assessments were based on a single interview and necessarily include measurement error. Analysis of data from multiple waves would permit tests of longitudinal stability and the partial control of measurement error. It is unclear whether item saliency changes over time. Although one might expect item thresholds to change, it is possible for the factor loadings to remain relatively stable. (4) Endorsement rates for several items were extremely low. This may result in unstable parameter estimates as well as artificially inflating the discrimination parameters. This was evident in the slope of the IRCs for ‘used more or longer than thought/planned’ among cocaine and hallucinogen users. Because discrimination is a function of the factor loadings and residual error, we have therefore presented 95% bootstrapping confidence intervals for the estimated factor loadings. (5) Our method permits the estimation of factor loadings that relate both the stem (use) and the probe (symptoms of dependence or abuse) items to the latent trait. However, the identification of this model relies on the assumption that the trichotomy of use into ‘none’ versus ‘six times or less’ versus ‘more than six times’ is unidimensional. We tested this assumption in monozygotic twin pairs and found no evidence for its violation. (6) Polysubstance abusers, by definition, do not use drugs in isolation. If, for example, there are users who abuse cannabis and cocaine simultaneously, then this may cause a minor tendency for IRC parameters of these substances to converge. Although we do not have data to show what proportion of the population do not use drugs in isolation, we have assumed that this will apply to only a small part of our sample. Finally, (7) there is the limitation of non-independent data. The current factor-analyses assumed that members of twin pairs were independent. Failure to take account of statistical non-independence, or violations of conditional independence across twins, is not expected to change parameter estimates, but may alter their confidence intervals. Exploratory factor-analyses are descriptive and not inferential, and therefore changes in the precision of the parameter estimates are unlikely to bias our results significantly. Nevertheless, we re-ran our analyses and included a twin-pair correlation between the latent factors, which varied according to zygotism as well as item-specific twin-pair correlations. There were no marked differences in our results.

**CONCLUSION**

Results from the factor-analysis reveal that criteria for abuse and dependence across the drug classes do not assess independent dimensions of drug-related problems. Instead, these items appear to index a single unidimensional continuum of risk. Inspection of the IRCs indicates
that many of the individual criteria do not assess the same levels of difficulty and discrimination across the six major illicit drug classes. Despite the unidimensionality, summing criteria within a substance may be inefficient, as individual criteria are likely to be redundant in terms of difficulty and assess varying levels of discrimination. Compared to other drugs, cocaine usage is associated with more detrimental effects, whereas the effects of cannabis and hallucinogens appear more benign and require greater levels of risk before the emergence of symptoms. Based on a large population-based sample of male twins, our results and the observed degrees of specificity are remarkably consistent with what is known about basic pharmacology of the drug groups in this study.

Acknowledgements

This project was supported by NIH grants DA-11287, MH/AA/DA-49492, MH-01458, AA-00236 and NHMRC Sidney Sax Postdoctoral Fellowship. The authors thank Indrani Ray for database assistance. We thank Dr Linda Corey for assistance with the ascertainment of twins from the Virginia Twins Registry, now part of the Mid-Atlantic Twin Registry (MATR), directed by Dr Judy Silberg. The registry has received support from NIH, the Carman Trust and the W. M. Keck, John Templeton and Robert Wood Johnson Foundations.

References


