Revising the Cannabis Use Disorders Identification Test (CUDIT) by means of Item Response Theory

BEATRICE ANNAHEIM,¹ THOMAS J. SCOTTO² & GERHARD GMEL^{1, 3}

1 Swiss Institute for the Prevention of Alcohol and Drug Problems, Lausanne, Switzerland

2 Department of Government, University of Essex, Colchester, UK

3 Alcohol Treatment Centre, Lausanne University Hospital, Switzerland

Key words

cannabis, CUDIT, screening, cannabis abuse, Item Response Theory

Correspondence

Beatrice Annaheim, Swiss Institute for the Prevention of Alcohol and Drug Problems
P.O. Box 870, CH-1001 Lausanne, Switzerland.
Telephone (+41) 21 321 29 94
Fax (+41) 21 321 29 40
Email: bannaheim@sfa-ispa.ch

Received 16 December 2008;
revised 26 April 2009;
accepted 29 May 2009

Abstract

Cannabis use among adolescents and young adults has become a major public health challenge. Several European countries are currently developing short screening instruments to identify 'problematic' forms of cannabis use in general population surveys. One such instrument is the Cannabis Use Disorders Identification Test (CUDIT), a 10-item questionnaire based on the Alcohol Use Disorders Identification Test. Previous research found that some CUDIT items did not perform well psychometrically. In the interests of improving the psychometric properties of the CUDIT, this study replaces the poorly performing items with new items that specifically address cannabis use.

Analyses are based on a sub-sample of 558 recent cannabis users from a representative population sample of 5722 individuals (aged 13–32) who were surveyed in the 2007 Swiss Cannabis Monitoring Study. Four new items were added to the original CUDIT. Psychometric properties of all 14 items, as well as the dimensionality of the supplemented CUDIT were then examined using Item Response Theory.

Results indicate the unidimensionality of CUDIT and an improvement in its psychometric performance when three original items (*usual hours being stoned; injuries; guilt*) are replaced by new ones (*motives for using cannabis; missing out leisure time activities; difficulties at work/school*). However, improvements were limited to cannabis users with a high problem score. For epidemiological purposes, any further revision of CUDIT should therefore include a greater number of 'easier' items. Copyright © 2010 John Wiley & Sons, Ltd.

Introduction

After alcohol and tobacco, cannabis is the most popular psychoactive substance in Europe (Vicente *et al.*, 2008). In a general population telephone survey on cannabis use in Switzerland, 43.5% of adolescents and young adults (aged 13–29) reported having used cannabis at least once

in their lifetime; 11.2% used it during the past six months, 4.1% on a weekly basis, and 0.9% on a daily basis (Arbeitsgruppe Cannabismonitoring, 2008). In a recent school survey, 9.8% of Swiss 14- and 15-year olds reported cannabis use during the last 30 days (Schmid *et al.*, 2008).

Although a majority of these users do not encounter any clinical or social problems, there is a growing public

health need to identify, in general population surveys, those who *do* have problems related to their cannabis use (Beck and Legleye, 2008). Intensive cannabis use can be considered as an indicator of problematic cannabis use. According to the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) (Vicente *et al.*, 2008), there are an estimated three million adults in Europe (1%) using cannabis daily or on an almost daily basis. Precise information on the prevalence of cannabis abuse, dependence and harmful use in Europe is limited. A recent review of surveys in Europe found that 0.3–2.9% of adults met criteria for ‘drug dependence’, which includes all illicit drugs (Vicente *et al.*, 2008). For several years, EMCDDA has reported an increasing number of people seeking treatment for cannabis use. However, reasons for this are multifaceted and might also be due to an expansion of treatment availability and an increased number of referrals for treatment by criminal justice officials (Montanari *et al.*, 2008).

In the scientific community, there is agreement on the possible negative consequences of cannabis use: beside its harmful effects on the respiratory system when smoked (Aldington *et al.*, 2007), there is evidence of its impairing effects on cognitive functioning and concentration as well as its detrimental effects on the course of schizophrenia and other mental disorders (e.g. Cohen *et al.*, 2008; Hall and Lynskey, 2009). However, there is no officially accepted definition of *problematic cannabis* use. At the same time, several easy-to-administer, short instruments exist which identify those cannabis users who encounter problems of any kind related to their use (e.g. Piontek *et al.*, 2008). However, there is no established screening instrument for problematic cannabis use as yet.

The Cannabis Use Disorders Identification Test (CUDIT: Adamson and Sellman, 2003, see Table 1) is a 10-item instrument that was constructed by adapting the Alcohol Use Disorders Identification Test (AUDIT: Saunders *et al.*, 1993). CUDIT’s performance was examined in clinical settings in New Zealand and Australia (Adamson and Sellman, 2003; Adamson *et al.*, 2008) as well as in a Swiss population survey (Annaheim *et al.*, 2008). In each context, some items performed weakly. For example, Item 9 (*injuries*) and Item 2 (*usual hours being stoned*) were endorsed only rarely, showed low scale correlations, or did not increase Cronbach’s α when included. These items, as derived from the alcohol screening test, proved unsuitable for the detection of problematic cannabis use.

When examining the psychometrics of CUDIT items, the aforementioned studies pragmatically assumed a one-dimensional construct, without empirically testing its

dimensionality. In contrast, the dimensionality of AUDIT has been examined in several studies and has produced mixed findings. Theoretically, the latter is a tridimensional construct (amount and frequency of use, dependence and harm), but in practice is often used as a unidimensional screening instrument. Empirical evidence points to mostly two-factor (drinking behaviour and adverse consequences, e.g. Bergman and Källmén, 2002; Kim *et al.*, 2008; Shields *et al.*, 2004), but also one-factor (Carey *et al.*, 2003) and four-factor models (frequency, quantity of use, binge-drinking and problems: Gmel *et al.*, 2001).

While originally designed for clinical application to adults, AUDIT was also found to be suitable for epidemiological research and for adolescents (e.g. Fleming, 1996; Ivis *et al.*, 2000; Reinert and Allen, 2007). In contrast to epidemiological population surveys, the symptoms of cannabis abuse and dependence, based on the international classifications of disease [Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition (DSM-IV) or International Classification of Diseases – 10th Revision (ICD-10)], can be assessed in clinical settings, instead of simply screening for cannabis-related problems. AUDIT has been validated in a multitude of different studies, which mostly undertake a comparison between it and DSM criteria of alcohol abuse and dependence (Reinert and Allen, 2002; Reinert and Allen, 2007). Originally, however, AUDIT was designed to screen for a broader concept of problematic alcohol use, including hazardous and harmful patterns of use prior to the onset of actual dependence or abuse (Saunders *et al.*, 1993).

The present study develops this point further. Its aim is to use CUDIT as a basis for the development of an instrument that can be used at an early stage to screen for a broader concept of general problematic patterns of cannabis use. For epidemiological research and population surveys, this is more appropriate than exclusively focusing on cannabis abuse and dependence, as the proper diagnosis of cannabis abuse and dependence requires trained clinical staff [American Psychiatric Association (APA), 1995], which is not a feature of population surveys.

As there is no generally accepted definition of *problematic cannabis* use to date, the present study adopts the working definition proposed by the EMCDDA. This definition considers problematic cannabis use as a consumption pattern that leads to negative social or health-related consequences for the user himself or for his environment (Beck and Legleye, 2008).

Traditionally, the psychometrics of screening or diagnostic instruments are assessed using classical test theory. However, in recent years, the use of Item Response Theory

Table 1 The Cannabis Use Disorders Identification Test (CUDIT), supplemented; frequencies and imputed missing values

Have you used any cannabis over the past six months? If YES, please answer the following questions about your cannabis use. Please tick the box that is most correct for you in relation to your cannabis use *over the past six months*.

	Percentage	Imputed missing values
1. How often do you use cannabis?		
1 monthly or less	47.3	
2 2–4 times a month	20.8	
3 2–3 times a week	16.1	
4 4 or more times a week	15.8	0
2. How many hours were you 'stoned' on a typical day when you had been using cannabis?		
0 1 or 2	71.9	
1 3 or 4	22.8	
2 5 or 6	4.7	
3 7 to 9	0.2	
4 10 or more	0.5	31
3. How often were you 'stoned' for six or more hours?		
0 never	55.0	
1 less than monthly	26.0	
2 monthly	11.8	
3 weekly	6.5	
4 daily or almost daily	0.7	15
4. How often during the past six months did you find that you were not able to stop using cannabis (once you had started)?		
0 never	83.3	
1 less than monthly	7.5	
2 monthly	3.8	
3 weekly	2.9	
4 daily or almost daily	2.5	7
5. How often during the past six months did you fail to do what was normally expected from you because of using cannabis?		
0 never	77.4	
1 less than monthly	16.7	
2 monthly	3.2	
3 weekly	2.3	
4 daily or almost daily	0.4	9
6. How often during the past six months did you need to use cannabis in the morning to get yourself going after a heavy session of using cannabis?		
0 never	87.5	
1 less than monthly	6.3	
2 monthly	2.3	
3 weekly	3.0	
4 daily or almost daily	0.9	6
7. How often during the past six months did you have a feeling of guilt or remorse after using cannabis?		
0 never	77.8	
1 less than monthly	12.0	
2 monthly	5.7	
3 weekly	2.7	
4 daily or almost daily	1.8	7

Table 1 *Continued*

Have you used any cannabis over the past six months? If YES, please answer the following questions about your cannabis use.

Please tick the box that is most correct for you in relation to your cannabis use *over the past six months*.

	Percentage	Imputed missing values
8. How often during the past six months have you had a problem with your memory or concentration after using cannabis?		
0 never	59.5	
1 less than monthly	21.1	
2 monthly	9.1	
3 weekly	7.7	
4 daily or almost daily	2.5	9
9. Have you or someone else been injured as a result of your use of cannabis over the past six months?		
0 no	94.6	
4 yes	5.4	4
10. Has a relative, friend or a doctor or other health worker been concerned about your use of cannabis over the past six months?		
0 no	79.0	
4 yes	21.0	4
11. During the past six months have you neglected your social environment, that is, your family, partner, friends, etc., because of using cannabis?		
0 no	90.9	
2 possibly	5.6	
4 yes	3.6	4
12. How often during the past six months have you refrained from taking part in leisure time activities that you originally wanted to do, e.g. going out, sports, hobbies, etc., because of using cannabis?		
0 never	79.4	
1 less than monthly	11.6	
2 monthly	4.8	
3 weekly	3.8	
4 daily or almost daily	0.4	6
13. How often during the past six months have you had difficulties at work or school, because of using cannabis?		
0 never	90.1	
1 less than monthly	5.9	
2 monthly	2.7	
3 weekly	0.9	
4 daily or almost daily	0.4	5
14. Which of the following statements best fits your personal situation?		
0 'I smoke cannabis for fun, because it's something special'	89.1	
4 'I smoke cannabis out of habit, because it's part of my daily life'	10.9	34

Note: $n = 558$; Items 1–10 see Adamson and Sellmann (2003); Items 11–14 were added by the authors.

(IRT) has grown in popularity (Embretson and Reise, 2000). IRT has several advantages over classical test theory in assessing self-reported health outcomes, e.g. the estimation of item difficulty according to the position of an individual on a latent trait variable and an adequate treatment of ordinal scaled and dichotomous items (Embretson and Reise, 2000; Hays *et al.*, 2000). In a number of recent studies, IRT was used to examine the diagnostic criteria for alcohol (e.g. Saha *et al.*, 2007) or for cannabis abuse and dependence (e.g. Hartman *et al.*, 2008; Langenbucher *et al.*, 2004; Martin *et al.*, 2006; Teesson *et al.*, 2002). In contrast, IRT has yet to be applied in the assessment of the psychometric performance of CUDIT.

Given this situation, the aim of the present paper was to explore the structure (dimensionality) of CUDIT and the performance of its items in a Swiss population sample of adolescents and young adults by means of IRT. Concretely, the objective was to develop CUDIT further by suggesting the inclusion of new better-performing items that are better suited to the study of problematic cannabis use. To this end, the following questions were addressed:

- What is the dimensionality of CUDIT?
- What are the psychometric properties and characteristics of its items?
- Are there alternative items with a better psychometric performance that could be included in a revised version of CUDIT?

Methods

Study design and sample

Analyses were based on data from the *Swiss Cannabis Monitoring Study*, which surveyed a representative general population sample of adolescents and young adults, aged 13 to 32 (Arbeitsgruppe Cannabismonitoring, 2008). This longitudinal survey mainly collected information about the amount of cannabis used, the resulting social and health consequences, as well as general attitudes towards cannabis and its legal status.

The first wave involved 5025 adolescents and young adults. Data analysed in the present study come from the second wave, conducted in 2007 and comprising 5722 participants. From the first wave, 3111 respondents were panelled and re-interviewed; the remaining 2611 were recruited in 2007 to compensate for panel attrition. Response rate in the second wave was 61.9%. Adolescents aged 13 to 18 were oversampled to enable us to make representative statements for this age-group. Two-stage random sampling was used (1. households, 2. person

within a household), stratified by region. Participants were interviewed by means of computer-aided telephone interviews.

In the second wave sample, 558 respondents were recent cannabis users, i.e. they used cannabis at least once during the six months preceding the interview. This group participated in the CUDIT part of the survey. Within this sub-sample ($n = 558$), 67.2% were male and 32.8% female; 62.4% were adolescents (13–18 years old) and 37.6% were 19 to 32 years old; mean age was 18.3 years. The sub-sample consisted of 12.2% daily or almost daily users; 28.7% used cannabis on a weekly basis (i.e. one to five times per week), 33.9% on a monthly basis, and 25.2% had used cannabis at least once during the last six months, but not on a monthly basis.

Measures

The original CUDIT consists of 10 items. It was supplemented by four additional items (see Table 1), as the first wave analyses had indicated the poor performance of several of the original CUDIT items (Annaheim *et al.*, 2008). The four new items were constructed based on (a) qualitative interviews, broaching the issue of negative consequences of intensive cannabis use with (former) users and with experts from the cannabis treatment field; (b) analyses of existing items from other available screening instruments; and (c) theoretical considerations. (Transcripts and synthesis of the qualitative interviews, a complete list of considered screening instruments, as well as notes on theoretical conclusions that led to the selection of the four new items are available on request from the authors.)

Analyses

To begin, we calculated frequencies over all 14 CUDIT items. Missing values on these items were imputed by means of Markov Chain–Monte–Carlo estimates (Congdon, 2003), available in the PRELIS software package (Jöreskog and Sörbom, 2002). Imputation replaced incomplete observations on some items with complete information, based on estimations of the values drawn from complete cases with the same answer patterns on all items but those missing.

IRT is based on the assumption of an underlying latent trait variable (in our case: problematic cannabis use), which is normally distributed and continuous. The latent trait variable, also referred to as latent factor, is estimated by manifest factor indicators, i.e. the items contained in a screening test, for example (Embretson and Reise, 2000). Adequacy of IRT-based measurements depends on two

interrelated preconditions: (1) unidimensionality, and (2) local independence. Unidimensionality means that a single latent trait variable is sufficient to explain the common variance among item responses, while local independence means that examinees' item responses depend only on their latent trait level, and not on their responses to other test items (Embretson and Reise, 2000).

To test these preconditions, we conducted exploratory and confirmatory factor analyses to assess the dimensionality of CUDIT. In addition, we examined model fit using the comparative fit index (CFI), the root mean square error of approximation (RMSEA), Chi-square statistics, and classical Cronbach's α . The CFI reports how much of the observed co-variance the model accounts for, relative to a naïve model, and model fit is judged as good if its value is greater than 0.90 (e.g. Hu and Bentler, 1999). RMSEA, which can be thought of as the average discrepancy between the observed and model-implied correlations, weighted by the degrees of freedom, should be lower than 0.08 (e.g. Browne and Cudeck, 1992; Hu and Bentler, 1999). Chi-square should not be statistically significant (i.e. $p < 0.05$), as this would indicate that there is a significant difference between the estimated model and the observed data. Finally, internal consistency, measured by Cronbach's α , is good if the statistic reaches or exceeds 0.80 (e.g. Cortina, 1993). However, such indices merely *hint* at the model fit. These cannot be treated as reliable decision-making tools, given that the topic of model fit in IRT is currently the subject of research and that there are no generally accepted fit indices (Embretson and Reise, 2000; Reeve *et al.*, 2007). In addition, to test for local independence, we computed residual item correlations from a confirmatory factor analysis, and considered residual correlations greater than 0.20 (Reeve *et al.*, 2007) as a possible indication of local dependence.

Given that CUDIT consists of dichotomous and ordered categorical factor indicators (items), and that our interest was item discrimination and location on the latent trait while guessing a 'correct' answer was irrelevant to our purpose, we employed the two-parameter Samejima's graded response model (Embretson and Reise, 2000; Rost, 1988; Samejima, 1996). In the Samejima's graded response model, each item is described by a slope or discrimination parameter (α) and by between-category threshold or location parameters (β). There is one α -parameter for each dichotomous or polytomous item. The higher the α parameter, the better the ability of the item to discriminate between individuals with different values on the latent trait. However, the item location or the *threshold* parameters (β) represent the points along the latent trait (x-axis) at which the probability of the

response lying in category k_i or higher exceeds 0.50 (k = number of response categories). There are $k - 1$ β -thresholds for each item. Higher β values indicate more difficult item response categories. The wide spacing of threshold parameters (β), ideally from negative to positive, is generally an advantage. An item with a wider range of thresholds is more representative than one with a narrower range, because the latter discriminates only for individuals with a certain position on the latent trait (Baker, 2001; Embretson and Reise, 2000).

Exploratory and confirmatory factor analyses, as well as tests of model fit, were calculated using Mplus software (Muthén and Muthén, 2008). In line with IRT, the metric of the factor was defined by fixing the factor variance, rather than the first factor loading, at one (Muthén *et al.*, 1997; Muthén and Muthén, 2008). Graphs were drawn in Microsoft Excel, based on Mplus output, while IRT α - and β -parameters and standard errors were calculated using Multilog (Thissen *et al.*, 2008).

To visualize the performance of each item, *Category Response Curves* (CRC) and *Items Information Curves* (IIC) were plotted. CRC show the probabilities of individuals endorsing the item response categories across a range of different latent trait levels. When plotting IRT graphs, the standardised latent trait values [mean = 0; standard deviation (SD) = 1] are represented on the x-axis, while the corresponding probabilities of endorsing different response categories are displayed on the y-axis. Each response category is represented by a separate curve (CRC). The steeper a curve is, the greater the ability of the response category to discriminate between individuals with different values on the latent trait. Yet, the further to the right of the x-axis the curve lies, the more difficult the response category; in other words, fewer respondents will endorse the category (Baker, 2001; Embretson and Reise, 2000).

For IIC, the amount of psychometric information (i.e. precision of estimation) provided by an item is plotted against the latent trait. In contrast to classical test theory where an item has only one reliability, in IRT an item provides more or less information depending on the positions of the respondent on the latent trait. The peak of the IIC represents the level on the latent trait where the item provides the most information. Higher information values indicate better psychometric performance of the item. In addition, the more broadly spread an IIC, the greater its ability to provide information. For the Test Information Curve (TIC), values of all items are summed and plotted in one single curve for all possible points on the latent trait (Baker, 2001; Embretson and Reise, 2000; Hays *et al.*, 2000).

Results

Table 1 reports the distribution of answers across the original CUDIT items (Items 1–10) and the four new ones, as well as the number of missing values that were imputed. On Item 1 (*frequency of use*), all individuals scored at least one point, as CUDIT was administered exclusively to cannabis users. Items 9 (*injuries*), 11 (*neglected social environment*), 13 (*difficulties at work/school*), 14 (*motives*), and 4 (*not able to stop*) were rarely endorsed, i.e. less than 20% of the subjects scored at least one point on these items. Finding an adequate response category appeared difficult for Items 14, 2 (*usual hours being stoned*) and 3 (*stoned for six or more hours*), as these show more than 10 missing values. The median missing values per item is 6.5. All items are right-skewed, which means that the higher the category (standing for more problematic use) was, the less it was endorsed.

Dimensionality of the CUDIT and local independence

Exploratory factor analyses revealed high eigenvalue ratio of the first (eigenvalue = 6.24; CUDIT, supplemented) to the second (1.43) and third (1.22) factor, suggesting a one-factor model (Montanelli and Humphreys, 1976). In that unidimensional model, residual correlations – as indicators of local independence – were rather low, with the

exception of Items 2 and 3 which showed a residual correlation of 0.291 (Appendix, Table A1).

However, model fit was moderate for the unidimensional model (CFI = 0.913, RMSEA = 0.079, Chi-square $p < 0.000$, Cronbach's $\alpha = 0.805$; CUDIT, supplemented), but improved when switching from the unidimensional to the bidimensional, and from the bidimensional to the tridimensional model (Appendix, Table A2). Nevertheless, the empirically extracted factors did not make much sense when it came to the content of the items: either items were grouped together in ways which did not make much sense (e.g. Items 1, 5, 8, 12 on factor one, and Items 3, 4, 6, 10, 13, 14 on factor two in the bidimensional model of CUDIT, revised); or Items 2 and 3 (excessive use) were integrated as one factor, which is logical but does not lead to a practicable factor since Item 2 performs so badly that it has to be deleted.

Finally, a one-dimensional approach was pursued for the subsequent IRT analyses based on the following considerations: (a) high first to second eigenvalue ratios; (b) low local dependence in the unidimensional model; and (c) the fact that empirically sound multiple factors were not logical when it came to the content of the items.

Psychometric properties and items' characteristics

The item parameters α and β , estimated according to the Graded Response Model, are reported in Table 2. Item 1,

Table 2 Slope (α) and category threshold parameters (β) estimated by the Graded Response Model; CUDIT, supplemented

Items	Slope α	Threshold β_1	Threshold β_2	Threshold β_3	Threshold β_4
1 frequency of use	2.11 (0.19)	–	–0.10 (0.07)	0.61 (0.08)	1.31 (0.11)
2 usual hours being stoned	0.77 (0.15)	1.37 (0.27)	4.02 (0.76)	6.72 (1.47)	7.10 (1.55)
3 stoned for 6 or more hours	1.67 (0.17)	0.17 (0.08)	1.26 (0.12)	2.13 (0.19)	3.68 (0.50)
4 not able to stop	2.46 (0.32)	1.19 (0.08)	1.62 (0.12)	1.95 (0.15)	2.39 (0.21)
5 failed to do what expected	1.54 (0.21)	1.11 (0.12)	2.35 (0.24)	2.97 (0.33)	4.37 (0.74)
6 morning use	1.82 (0.28)	1.59 (0.15)	2.12 (0.21)	2.45 (0.25)	3.35 (0.45)
7 guilt/remorse	1.15 (0.17)	1.33 (0.18)	2.25 (0.28)	3.10 (0.40)	3.97 (0.55)
8 memory/concentration problems	2.17 (0.20)	0.30 (0.07)	1.08 (0.09)	1.59 (0.12)	2.50 (0.22)
9 injuries	0.72 (0.29)	4.28 (1.53)	–	–	–
10 concerned others	1.27 (0.20)	1.35 (0.18)	–	–	–
11 neglected social environment	1.09 (0.23)	2.51 (0.43)	3.50 (0.61)	–	–
12 missed out leisure time activities	1.44 (0.20)	1.24 (0.13)	2.05 (0.21)	2.73 (0.30)	4.57 (0.81)
13 difficulties at school or work	1.90 (0.35)	1.74 (0.16)	2.37 (0.24)	3.07 (0.39)	3.78 (0.64)
14 motives for using cannabis	2.25 (0.37)	1.56 (0.14)	–	–	–

Note: Scores in parentheses: standard error values; $n = 558$; estimator = marginal maximum likelihood (MML; Multilog). The higher the slope parameter α , the better the item discriminates per latent trait. Between-category threshold parameters β represent the points along the latent trait at which the probability of responding to a certain category passes 0.50. The greater the variety of threshold parameters β (negative to positive), the more descriptive the item is.

followed by Item 8 and then Item 3 show high discrimination and are rather easy ($\alpha > 1.60$ /first β -thresholds ≤ 0.30), while Item 9, followed by items 2, 11, 7 discriminate less well ($\alpha < 1.20$). Items 9, 11, 13, 6 and 14 are relatively difficult (first β thresholds > 1.50).

These findings are visualised by CRC in Figure 1.

In Figure 2, IIC are plotted for all 14 CUDIT items. Item 9, followed by Items 2, 11 and 7 provide the least psychometric information (lowest peaks).

Revising the CUDIT

The analyses, reported above as IRT parameters, CRC, IIC, test for local dependence, frequency distributions, show that original Items 9, 2, and 7 perform poorly. We suggest replacing these with the following new items: 12 (*missed out leisure time activities*), 14 (*motives*) and 13 (*problems at work/school*).

Compared to the original version, the revised CUDIT does not show an improved model fit when measured by RMSEA or chi-square (Appendix, Table A2). However, model fit is slightly improved when measured by CFI (0.935 versus 0.927), and internal consistency (Cronbach's α) also increased (from 0.759 to 0.800). For individuals that are above the mean on the latent trait, the revised version discriminates better and gives more psychometric information than the original, as shown by a steeper and higher TIC (Figure 3).

Discussion

Cannabis use has become a major public health concern in developed societies. In several European countries, work is under way to develop and validate adequate instruments to monitor intensive, dependent or problematic forms of cannabis use in population surveys (EMCDDA, 2007).

The aim of this paper was *not* to present an accomplished new screening instrument for problematic cannabis use, but rather to explore the psychometric properties of an existing instrument (CUDIT), and to make some suggestions for its improvement and further development. In this spirit, we conclude the following.

The results indicate that CUDIT could be treated as a unidimensional construct, measuring one latent variable which could be described as 'problematic cannabis use'. At the same time, the psychometric performance of some of its items was poor.

Item 9 (*injuries*) has shown a weak performance across all criteria (α and β parameters, IIC, etc.). This result confirms findings of earlier studies that used classical test theory and were carried out in a clinical population by the CUDIT authors (Adamson and Sellman, 2003) and by the authors of the present paper in a previous wave of the Swiss Cannabis Monitoring Study (Annaheim *et al.*, 2008). Injuries can either stem from accidents or from physical aggression. Therefore, our results are not

Figure 1 CRC; CUDIT, supplemented (items 1–14: left to right, top–down). Note: $n = 558$; estimator = maximum likelihood with robust standard errors (MLR; Mplus). CRCs indicate the probability of endorsing the item at a given response category across locations on the latent trait. Well performing items have CRCs with sharper and more differentiated peaks across response categories.

Figure 2 IIC; CUDIT, supplemented. Note: $n = 558$; estimator = maximum likelihood with robust standard errors (MLR; Mplus). IIC show the information given by the single items along the latent trait. The higher an IIC is the more information is given.

Figure 3 TIC; comparison of CUDIT, original and CUDIT, revised. Note: $n = 558$; estimator = maximum likelihood with robust standard errors (MLR; Mplus); TIC show the information given by a whole test along the latent trait. The higher an TIC is, the more information the test delivers.

surprising, as alcohol use is more clearly linked to aggression than cannabis use (e.g. Gillet *et al.*, 2001; Hoaken and Stewart, 2003). In addition, even if cannabis is used in dangerous situations like in traffic, the chance of being involved in an accident *and* of someone consequently being injured is very small. We would therefore suggest a replacement of the Item 9 (*injuries*) by Item 13, which assesses the effective difficulties at work or school, as these are often experienced by intensive cannabis users (e.g. Hall *et al.*, 2001; Thomas *et al.*, 2006).

The poor performance of Item 2 (*usual hours being stoned*) also confirms results from a previous CUDIT analysis (Annaheim *et al.*, 2008) and from other studies showing that quantity of use is commonly a poor predictor of problematic consumption patterns (namely, dependence), while frequency of use (Chen *et al.*, 1997) and episodic excessive use (so called 'binge', Item 3) are more predictive. Our results show that only a small proportion (0.7%) of cannabis users scored more than two points on Item 2. Other researchers have faced the same problem with this item and tried to improve its performance by grouping response categories into smaller entities (<one hour in stead of 'one or two' hours, 'one or two hours' in stead of 'three or four' hours, etc., Adamson *et al.*, 2008). As the number of hours being stoned is already well assessed by Item 3, we suggest replacing Item 2 by Item 14 (*motives for using cannabis*), since motivation is crucial when defining problematic patterns of use (Brodbeck *et al.*, 2007).

Although Item 7 of CUDIT (*guilt/remorse*) is the same as Item 7 of AUDIT, it is questionable whether one can have a feeling of guilt on a 'daily', 'weekly', or 'monthly' basis; feelings of guilt seem to be more generalized emotional states. In addition, they are highly influenced by societal climate, social norms and the social environment of a cannabis user. The content of Item 7, therefore, might explain its weak performance in our analyses. We suggest replacing it by Item 12 (*missed out leisure time activities*), which assesses an effective constraint on life quality due to cannabis use as perceived by the users themselves.

The fourth supplementary CUDIT item (Item 11: *neglected social environment*) did not perform better than any of the seven remaining original items. Although users in the previous qualitative interviews had mentioned neglect of family and friends as an important negative aspect of intensive cannabis use, Item 11 was rarely endorsed by individuals in the quantitative study. This might be due to the fact that the item was too 'difficult', that it is hard to admit to neglecting one's social environment and, as observed for Item 7, response categories in

the form of 'daily', 'weekly', etc. might have led to problems.

Our study, like others, faces some limitations. First, CUDIT was not compared to a reference 'gold' standard. Cannabis abuse, dependence or harmful use, according to DSM or ICD criteria might be possible 'gold standards' to which the CUDIT could be compared (see Adamson and Sellman, 2003). However, in our opinion and in accordance with AUDIT (Saunders *et al.*, 1993), CUDIT, when used for epidemiological purposes, is not intended to solely measure cannabis abuse or dependence, but a broader concept of *problematic cannabis use* prior to the onset of dependence or harm, for which no reference standard currently exists. Second, we have simply added four additional items to the original CUDIT instead of an 'item pool'; however, these items were tested in a representative population sample. Third, we decided to treat CUDIT as a unidimensional model, although model fit could have been improved by adding more factors. However, the empirically extracted, better fitting models did not make sense when it came to item content. In addition, there was rather low local dependence in the unidimensional model. Our decision to treat CUDIT as a unidimensional scale is confirmed by the literature, which states that if model assumptions are supported by the data, then strict adherence to model fit statistics is not vital, and 'essential' unidimensionality (McDonald, 1981) is sufficient for satisfying the IRT assumption of unidimensionality. Furthermore, over-reliance on fit statistics are not desirable given that universally accepted fit statistics for IRT do not yet exist (Embretson and Reise, 2000; Hays *et al.*, 2000; Reeve *et al.*, 2007).

Seeing that for our population-based sample, all items – except Item 1 – have β -threshold parameters greater than zero (i.e. are difficult), a further revised version of CUDIT would need more 'easier' questions, i.e. items that discriminate better for cannabis users who have a below-average problem score. As we have shown by the TIC (Figure 3), the improvement of the revised CUDIT over the original concerns mainly individuals who are above the mean on the latent trait. Ideally, though, an improvement in measuring precision for epidemiological purposes should also be attained for individuals whose cannabis use is less problematic.

Acknowledgements

The Swiss Cannabis Monitoring Study is commissioned and financed by the Swiss Federal Office of Public Health (SFOPH; contract number 01.001316). The following institutions and researchers contribute to the study: Swiss Institute for the Prevention of Alcohol and Drug Problems (SIPA),

Lausanne (G. Gmel, B. Annaheim); Research Institute for Public Health and Addiction (ISGF/RIPHA), Zurich (A. Uchtenhagen, M. Schaub); Institute of Social and Preventive Medicine (IUMSP), Lausanne (F. Dubois-Arber, S. Arnaud, J. P. Gervasoni); Institute of Criminology and Criminal Law (ICDP) (M. Killias, J. Vuille); Lausanne; Institut fuer Begleit- und Sozialforschung (IBSF), Zurich (M. Mueller). The SFOPH approved the publication of the current investigation conducted as part of the Swiss Cannabis Monitoring Study.

Declaration of interest statement

The authors have no competing interests.

References

- Adamson S.J., Sellman J.D. (2003) A prototype screening instrument for cannabis use disorder: The Cannabis Use Disorders Identification Test (CUDIT) in an alcohol-dependent clinical sample. *Drug and Alcohol Review*, **22**, 309–315.
- Adamson S.J., Sellman J.D., Kay-Lambkin F., Baker A., Lewin T., Thornton L., Kelly B. (2008) A revised screening measure for cannabis misuse: The Cannabis Use Disorders Identification Test (CUDIT), *XXIX International Congress of Psychology*, Berlin.
- Aldington S., Williams M., Nowitz M., Weatherall M., Pritchard A., McNaughton A., Robinson G., Beasley R. (2007) Effects of cannabis on pulmonary structure, function and symptoms. *Thorax*, **62**, 1058–1063, DOI: 10.1136/thx.2006.077081
- American Psychiatric Association (APA) (1995) *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV*, APA.
- Annaheim B., Rehm J., Gmel G. (2008) How to screen for problematic cannabis use in population surveys. An evaluation of the Cannabis Use Disorders Identification Test (CUDIT) in a Swiss sample of adolescents and young adults. *European Addiction Research*, **14**, 190–197, DOI: 10.1159/000141643
- Arbeitsgruppe Cannabismonitoring (2008) Veränderungen im Cannabiskonsum 2004 bis 2007 – Ergebnisse des Schweizerischen Cannabismonitorings. <http://www.bag.admin.ch/themen/drogen/00042/00643/00646/02287/index.html?lang=de> [15 October 2008].
- Baker F. (2001) *The Basics of Item Response Theory*. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland.
- Beck F., Legleye S. (2008) Measuring cannabis related problems and dependence at the population level. In: *A Cannabis Reader: Global Issues and Local Experiences* (eds European Monitoring Center for Drugs and Drug Addiction (EMCDDA)), EMCDDA.
- Bergman H., Källmén H. (2002) Alcohol use among Swedes and a psychometric evaluation of the alcohol use disorders identification test. *Alcohol and Alcoholism*, **37**, 245–251.
- Brodbeck J., Matter M., Page J., Moggi F. (2007) Motives for cannabis use as a moderator variable of distress among young adults. *Addictive Behaviors*, **32**, 1537–1545, DOI: 10.1016/j.addbeh.2006.11.012
- Browne M.W., Cudeck R. (1992) Alternative ways of assessing model fit. *Sociological Methods & Research*, **21**, 230–258, DOI: 10.1177/0049124192021002005
- Carey K.B., Carey M.P., Chandra P.S. (2003) Psychometric evaluation of the alcohol use disorders identification test and short drug abuse screening test with psychiatric patients in India. *Journal of Clinical Psychiatry*, **64**, 767–774.
- Chen K., Kandel D.B., Davies M. (1997) Relationships between frequency and quantity of marijuana use and last year proxy dependence among adolescents and adults in the United States. *Drug and Alcohol Dependence*, **46**, 53–67.
- Cohen M., Solowij N., Carr V. (2008) Cannabis, cannabinoids and schizophrenia: Integration of the evidence. *Australian and New Zealand Journal of Psychiatry*, **42**, 357–368, DOI: 10.1080/00048670801961156
- Congdon P. (2003) *Applied Bayesian Modelling*, John Wiley & Sons.
- Cortina J.M. (1993) What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, **78**, 98–104.
- Embretson S., Reise S. (2000) *Item Response Theory for Psychologists*, Erlbaum.
- European Monitoring Center for Drugs and Drug Addiction (EMCDDA) (2007) *Annual Report 2007: The State of the Drugs Problem in Europe*, EMCDDA.
- Fleming J.M. (1996) The epidemiology of alcohol use in Australian women: Findings from a national survey of women's drinking. *Addiction*, **91**, 1325–1334.
- Gillet C., Polard E., Mauduit N., Allain H. (2001) Passage à l'acte et substances psychoactives: Alcool, médicaments, drogues. *L'Encéphale*, **27**, 351–359.
- Gmel G., Heeb J.-L., Rehm J. (2001) Is frequency of drinking an indicator of problem drinking? A psychometric analysis of a modified version of the alcohol use disorders identification test in Switzerland. *Drug and Alcohol Dependence*, **64**, 151–163.
- Hall W.D., Degenhardt L.J., Lynskey M.T. (2001) *The Health and Psychological Effects of Cannabis Use*, Commonwealth of Australia.
- Hall W., Lynskey M. (2009) The challenges in developing a rational cannabis policy. *Current Opinion in Psychiatry*, **22**, 258–262, DOI: 10.1097/YCO.0b013e3283298f36
- Hartman C.A., Gelhorn H., Crowley T.J., Sakai J.T., Stallings M., Young S.E., Rhee S.H., Corley R., Hewitt J.K., Hopfer C.J. (2008) Item response theory analysis of DSM-IV cannabis abuse and dependence criteria in adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, **47**, 165–73, DOI: 10.1097/chi.0b013e31815cd9f2

- Hays R.D., Morales L.S., Reise S.P. (2000) Item response theory and health outcomes measurement in the 21st century. *Medical Care*, **38**, II28–II42.
- Hoaken P.N.S., Stewart S.H. (2003) Drugs of abuse and the elicitation of human aggressive behavior. *Addictive Behaviors*, **28**, 1533–1554.
- Hu L., Bentler P.M. (1999) Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, **6**, 1–55.
- Ivis F.J., Adlaf E.M., Rehm J. (2000) Incorporating the AUDIT into a general telephone survey: A methodological experiment. *Drug and Alcohol Dependence*, **60**, 97–104.
- Jöreskog K.G., Sörbom D. (2002) *PRELIS 2: User's Reference Guide*, SSI.
- Kim S.S., Gulick E.E., Nam K.A., Kim S.H. (2008) Psychometric properties of the alcohol use disorders identification test: A Korean version. *Archives of Psychiatric Nursing*, **22**, 190–199, DOI: 10.1016/j.apnu.2007.07.005
- Langenbucher J.W., Labouvie E., Martin C.S., Sanjuan P.M., Bavy L., Kirisci L., Chung T. (2004) An application of item response theory analysis to alcohol, cannabis, and cocaine criteria in DSM-IV. *Journal of Abnormal Psychology*, **113**, 72–80, DOI: 10.1037/0021-843X.113.1.72
- Martin C.S., Chung T., Kirisci L., Langenbucher J.W. (2006) Item response theory analysis of diagnostic criteria for alcohol and cannabis use disorders in adolescents: Implications for DSM-V. *Journal of Abnormal Psychology*, **115**, 807–814, DOI: 10.1037/0021-843X.115.4.807
- McDonald R.P. (1981) The dimensionality of test and items. *British Journal of Mathematical and Statistical Psychology*, **34**, 100–117.
- Montanari L., Taylor C., Griffiths P. (2008) Cannabis users in drug treatment in Europe: An analysis from treatment demand data. In: *A Cannabis Reader: Global Issues and Local Experiences* (eds EMCDDA), pp. 261–278, EMCDDA.
- Montanelli R., Humphreys L. (1976) Latent roots of random data correlation matrices with squared multiple correlations on the diagonal: A Monte Carlo study. *Psychometrika*, **41**, 341–348.
- Muthén B.O., du Toit S.H.C., Spisic D. (1997) Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. http://www.gseis.ucla.edu/faculty/muthen/articles/Article_075.pdf [27 November 2008].
- Muthén L.K., Muthén B.O. (2008) *Mplus: User's Guide 5.1*. <http://www.statmodel.com/ugexcerpts.shtml> [14 October 2008].
- Piontek D., Kraus L., Klempova D. (2008) Short scales to assess cannabis-related problems: A review of psychometric properties. *Substance Abuse Treatment, Prevention, and Policy*, **3**, 25, DOI: 10.1186/1747-597X-3-25
- Reeve B.B., Hays R.D., Bjorner J.B., Cook K.F., Crane P.K., Teresi J.A., Thissen D., Revicki D.A., Weiss D.J., Hambleton R.K., Liu H., Gershon R., Reise S.P., Lai J.-S., Cella D. (2007) Psychometric evaluation and calibration of health-related quality of life item banks: Plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical Care*, **45**, S22–S31, DOI: 10.1097/01.mlr.0000250483.85507.04
- Reinert D., Allen J.P. (2002) The Alcohol Use Disorders Identification Test (AUDIT): A review of recent research. *Alcoholism: Clinical and Experimental Research*, **26**, 272–279.
- Reinert D.F., Allen J.P. (2007) The alcohol use disorders identification test: An update of research findings. *Alcoholism: Clinical and Experimental Research*, **31**, 185–199, DOI: 10.1111/j.1530-0277.2006.00295.x
- Rost J. (1988) *Quantitative und qualitative probabilistische Testtheorie*, Hans Huber.
- Saha T.D., Stinson F.S., Grant B.F. (2007) The role of alcohol consumption in future classifications of alcohol use disorders. *Drug and Alcohol Dependence*, **89**, 82–92, DOI: 10.1016/j.drugalcdep.2006.12.003
- Samejima F. (1996) The graded response model. In: *Handbook of Modern Item Response Theory* (eds van der Linden W, Hambleton RK), pp. 85–100, Springer.
- Saunders J.B., Aasland O.G., Babor T.F., De la Fuente J.R., Grant M. (1993) Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption – II. *Addiction*, **88**, 791–804.
- Schmid H., Delgrande Jordan M., Kuntsche E., Kuendig H., Annaheim B. (2008) *Der Konsum psychoaktiver Substanzen von Schülerinnen und Schülern in der Schweiz – Ausgewählte Ergebnisse einer Studie, durchgeführt unter der Schirmherrschaft der Weltgesundheitsorganisation (WHO), Schweizerische Fachstelle für Alkohol- und andere Drogenprobleme (SFA)*.
- Shields A.L., Guttmanova K., Caruso J.C. (2004) An examination of the factor structure of the Alcohol Use Disorders Identification Test in two high-risk samples. *Substance Use and Misuse*, **39**, 1161–1182.
- Teesson M., Lynskey M., Manor B., Baillie A. (2002) The structure of cannabis dependence in the community. *Drug and Alcohol Dependence*, **68**, 255–262.
- Thissen D., Chen W.-H., Bock D. (2008) *Multilog 7*, Scientific Software International.
- Thomas G., Flight J., Richard K., Racine S. (2006) *Toward a Policy-Relevant Typology of Cannabis Use for Canada*, Health Canada, Canadian Executive Council on Addictions & Canadian Centre on Substance Abuse.
- Vicente J., Olszewski D., Matias J. (2008) Prevalence, patterns and trends of cannabis use among adults in Europe. In: *A Cannabis Reader: Global Issues and Local Experiences* (eds EMCDDA), pp. 3–28, EMCDDA.

Appendix

Table A1 Residual correlation and residual variance (*diagonal*); CUDIT, supplemented

	Supplemented CUDIT, 14 items													
	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Item 12	Item 13	Item 14
1 frequency of use	0.316	–	–	–	–	–	–	–	–	–	–	–	–	–
2 usual hours being stoned	–0.086	0.836	–	–	–	–	–	–	–	–	–	–	–	–
3 stoned for six or more hours	0.044	0.291	0.535	–	–	–	–	–	–	–	–	–	–	–
4 not able to stop	0.015	–0.140	–0.080	0.402	–	–	–	–	–	–	–	–	–	–
5 failed to do what expected	–0.174	–0.061	–0.057	–0.046	0.548	–	–	–	–	–	–	–	–	–
6 morning use	0.151	–0.095	0.017	0.086	–0.155	0.503	–	–	–	–	–	–	–	–
7 guilt/remorse	–0.093	–0.073	–0.098	0.136	–0.017	–0.154	0.703	–	–	–	–	–	–	–
8 memory problems	–0.046	–0.031	–0.012	–0.008	0.115	–0.109	0.014	0.419	–	–	–	–	–	–
9 injuries	–0.003	–0.091	–0.059	–0.166	–0.006	–0.129	0.045	0.090	0.895	–	–	–	–	–
10 concerned others	0.005	0.128	0.065	0.062	–0.135	–0.003	0.052	–0.094	0.042	0.652	–	–	–	–
11 neglected social environment	–0.089	–0.150	0.055	–0.013	0.149	–0.070	0.094	–0.013	0.154	–0.040	0.721	–	–	–
12 missed activities	–0.119	–0.033	0.016	–0.116	0.142	–0.147	0.045	0.062	0.083	–0.079	–0.039	0.613	–	–
13 difficulties at school/work	–0.127	–0.025	–0.135	–0.119	0.161	–0.160	0.002	0.097	0.038	0.011	0.086	0.098	0.494	–
14 motives using cannabis	0.144	–0.135	–0.133	0.064	–0.185	0.041	0.069	–0.065	–0.115	0.011	–0.028	–0.039	–0.079	0.416

Note: $n = 558$; estimator = weighted least square, mean- and variance-adjusted (WLSMV; Mplus); bold typeface: residual correlation > 0.20 , considered as local dependence (Reeve *et al.*, 2007).

Table A2 Model fit and internal consistency of CUDIT; unidimensional, bidimensional and tridimensional models according to exploratory factor analyses

	Unidimensional	Bidimensional	Tridimensional
<i>CUDIT, original: 10 items</i>			
Goodness of fit			
Comparative Fit Index (CFI)	0.927	0.957	0.988
Root Mean Square Error of Approximation (RMSEA)	0.090	0.077	0.047
Chi-square	137.714, df = 25, $p < 00.000$	86.235, df = 20, $p < 00.000$	33.739, df = 15, $p = 00.0037$
Internal consistency (Cronbach's α)			
First factor	0.759	0.719	0.639
Second factor	–	0.579	0.579
Third factor	–	–	0.535
<i>CUDIT, supplemented: 14 items</i>			
Goodness of fit			
Comparative Fit Index (CFI)	0.913	0.955	0.991
Root Mean Square Error of Approximation (RMSEA)	0.079	0.060	0.028
Chi-square	204.389, df = 46, $p < 0.000$	123.790, df = 41, $p < 0.000$	51.736, df = 36, $p = 0.0433$
Internal consistency (Cronbach's α)			
First factor	0.805	0.739	0.579
Second factor	–	0.685	0.701
Third factor	–	–	0.685
<i>CUDIT, revised: 10 items</i>			
Goodness of fit			
Comparative Fit Index (CFI)	0.935	0.992	–
Root Mean Square Error of Approximation (RMSEA)	0.091	0.034	–
Chi-square	140.876, df = 25, $p < 0.000$	34.641, df = 21, $p < 0.0309$	–
Internal consistency (Cronbach's α)			
First factor	0.800	0.703	–
Second factor	–	0.662	–

Note: $n = 558$; estimator = weighted least square, mean- and variance-adjusted (WLSMV; Mplus); CUDIT, original: bidimensional model: Items 1, 4, 5, 6, 7, 8, 9, 10 (factor one), Items 2, 3 (factor two); tridimensional model: Items 1, 4, 6, 10 (factor one), Items 2, 3 (factor two), Items 5, 7, 8, 9 (factor three); CUDIT, supplemented: bidimensional model: Items 1, 2, 3, 4, 6, 10, 14 (factor one), Items 5, 7, 8, 9, 11, 12, 13 (factor two); tridimensional model: Items 2, 3 (factor one), Items 1, 4, 6, 10, 14 (factor two), Items 5, 7, 8, 9, 11, 12, 13 (factor three); CUDIT, revised: unidimensional model: Items 1, 3, 4, 5, 6, 8, 10, 12, 13, 14; bidimensional model: Items 1, 5, 8, 12 (factor one), Items 3, 4, 6, 10, 13, 14 (factor two); tridimensional model: impossible to estimate the model, no convergence achieved (iterations = 200 000).