# The effect of sample selection on the distinction between alcohol abuse and dependence

**Martin Steppan[1], Daniela Piontek[1], and Ludwig Kraus[1,2]**

[1]IFT Institut für Therapieforschung, München, Germany
[2]Centre for Social Research on Alcohol and Drugs (SoRAD), Stockholm University, Stockholm, Sweden

## Abstract

**Aim:** The effect of sample selection on the dimensionality of DSM-IV alcohol and dependence (AUD) criteria was tested applying different methods.

**Sample:** Data from the 2006 German Epidemiological Survey of Substance Abuse (ESA) were used. A mixed-mode design was used (self-administered questionnaires and telephone interviews), and 7,912 individuals, aged 18 to 64 years, participated. The response rate was 45%. Alcohol abuse and dependence were assessed according to DSM-IV, based on the Munich Composite International Diagnostic Interview (M-CIDI). Inter-item correlations, Confirmatory Factor Analysis (CFA), and Latent Class Analysis (LCA) were applied to the total sample (unrestricted sample, URS) and a subsample of individuals with at least one endorsed criterion (restricted sample, RS). Latent Class Factor Analysis (LCFA) was performed using the RS, including covariates (age, sex, education).

**Findings:** The mean inter-item correlation was higher in the URS than in the RS. When individuals without criterion endorsement were excluded, factor analyses resulted in more dimensions. In the RS, LCA yielded an interaction between abuse, dependence and class membership. The LCFA identified two dimensions and five classes corresponding to abuse and dependence.

**Conclusions:** Sample selection has a critical effect on dimensionality analyses. When individuals who do not endorse a single criterion are excluded, the bi-axial factor structure of the DSM-IV (abuse and dependence) can be supported. However, there is also evidence that a further diagnostic category should be included or that the threshold for dependence should be lowered.

The recently published revision of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) combines abuse of and dependence on substances into one single disorder with graded severity, called substance use disorder (Regier, Kuhl, & Kupfer, 2013). This change was made in response to a number of studies indicating a one-dimensional structure of abuse and dependence criteria. These findings contradict the bi-axial concept of dependence and the consequences of use found in earlier versions of the DSM (Edwards & Gross, 1976).

A prominent critique of DSM-5 addresses the proposed abolishment of the abuse category (Edwards, 2012). According to Edwards, the decision overrelies on certain "American survey reports" at the population level, and neglects both clinical experience and a "significant body of epidemiological and anthropological research" (Edwards,

2012, p. 701). He argues that large-scale surveys might be insufficient for the distinction of relevant clinical subtypes.

Indeed, most studies assessing the dimensionality of abuse and dependence are based on large-scale general population samples. Sample restrictions in these studies usually involve the exclusion of people who have very low alcohol consumption (e.g., who have had fewer than 12 drinks during the previous year, or who have never drunk alcohol). Most of these studies found a one-dimensional structure using factor analytic approaches (Hasin & Beseler, 2009; Keyes, Krueger, Grant, & Hasin, 2011; Proudfoot, Baillie, & Teeson, 2006; Saha, Chou, & Grant, 2006; Slade, Grove, & Teeson, 2009). Applying latent class analytic techniques also lead to the assumption of one construct of graded severity (Baillie & Teeson, 2006; Bucholz et al., 1996; Lynskey & Agrawal, 2007).

If more restricted samples are used (heavy drinkers, clinical samples), more heterogeneous results are found (Borges et al., 2010; Gelhorn et al., 2008; Hardford & Muthén, 2001; Langenbucher et al., 2004; Martin, Chung, & Kirisci, 2006; Muthén, 1995; Nelson, Rehm, Üstün, Grant, & Chatterji, 1999; Proudfoot et al., 2006). The impact of sample selection was most impressively demonstrated by Nelson et al. (1999) in a study evaluating the factor structures of different DSM-IV substance use disorder criteria. Using the full sample of current substance users, the authors found a single dimension underlying abuse and dependence criteria. However, when excluding individuals at the lower (0 criteria) and higher (10 to 11 criteria) levels of severity, two dimensions were identified, reflecting the diagnostic categories of abuse and dependence. Regardless of sample restriction, the inclusion of covariates (Multivariate Indicators Multivariate Causes [MIMIC]) or the use of other psychological predictors (e.g., Neuroticism) also seem to produce more-dimensional results (Grant et al., 2007; Hardford & Muthén, 2001; Simons, Carey, & Wills, 2009). The inclusion of covariates that predict very low or very high criteria endorsement might have a similar (variance reducing) effect on analyses as the exclusion of these cases.

In sum, there is evidence that sample composition indeed matters when analysing the dimensionality of DSM-IV criteria. Although not conclusive, this evidence seems to indicate a tendency towards a one-dimensional structure when unrestricted general population samples are used, and towards more-dimensional structures with the use of more restricted samples. A reason for this effect could be that the inclusion of individuals not endorsing a single criterion ("zero class") leads to increased covariance between abuse and dependence. This is due to the fact that a zero class necessarily has a zero-variance on all of the observed criteria. Consequently, factor analyses and correlations overestimate the statistical overlap between abuse and dependence. One may argue that non-linear factor analytic techniques safeguard against this effect. Statistical software packages such as Mplus typically use tetrachoric correlations to perform non-linear factor analyses (Muthén & Muthén, 2001). However, considering that tetrachoric correlation is based on a four-cell table, an increasing percentage of a zero class will lead to a correlation approximating 1.

The present study aims at testing the effect of sample selection on the dimensionality of alcohol abuse and dependence, applying different methods. Using an empirical dataset from a general population survey, dimensionality analyses will be performed separately for a) the total sample (including the zero class) and b) a subsample of individuals with at least one endorsed criterion (excluding the zero class). It is expected that if abuse and dependence were separable constructs, both should be better identifiable when excluding the zero class. According to the literature, the effect of sample selection will be tested on the grouping of items (factor analysis) as well as the grouping of individuals (latent class analysis) and combinations of both (latent class factor analysis).

## Methods

### Sample

Data from the 2006 German Epidemiological Survey of Substance Abuse (ESA) were used (Kraus & Baumeister, 2008). A two-stage probability sampling design was used to select a random sample of non-institutionalized adults aged 18 to 64 years. In a first step, communities were selected proportional to population size. In a second step, individuals were randomly drawn from population registers, with younger age groups being oversampled and older age groups being undersampled. This disproportionate sampling strategy was used because the percentage of younger adults in the general population is considerably lower than the percentage of older individuals. A mixed-mode design was applied in the survey. All participants received a questionnaire by post. Those participants who did not wish to respond by post could answer the questionnaire by telephone. Mode-effect analyses generally revealed a high degree of comparability between the administration modes. Out of all eligible subjects, 7,912 individuals participated in the survey. Considering sample-irrelevant drop-outs (death, address unknown), the response rate was 45%. The ESA sample comprised 51.4% males; the prevalence of alcohol abuse and dependence was 2.6% and 3.9%, respectively. The sample did not perfectly represent the general population; in particular, people with lower educational levels and foreign nationalities were underrepresented.

### Measures

Alcohol abuse and dependence according to DSM-IV were assessed using the items of the 12-month version of the Munich Composite International Diagnostic Interview (M-CIDI) (Wittchen et al., 1995; Wittchen, 1994). The M-CIDI assesses the 11 alcohol abuse and dependence criteria (AUD) based on 18 questions (seven related to abuse, 11 to dependence). Alcohol dependence was diagnosed if at least three out of seven criteria were present in the last 12 months. For a diagnosis of alcohol abuse, at least one out of four criteria had to be met. Data on alcohol abuse and dependence criteria were available for 7,804 subjects (analytical sample). As covariates, age, gender and education were included. Educational attainment was measured on two items asking for the highest level of school completed and the highest professional qualification. Individual scores range from 1 (maximum secondary school and no completed professional qualification) to 7 (university degree) (Piontek, Kraus, Müller, & Pabst, 2010).

### Statistical Analyses

All analyses were performed on two samples. The unrestricted sample (URS) contained all individuals in the dataset ($N = 7,804$), including 6,104 respondents (78.2%) not fulfilling any criterion. The restricted sample (RS) contained only individuals who met at least one abuse or dependence criterion ($n = 1,700$). To test the effect of sampling (URS vs. RS) on the inter-item correlations, tetrachoric correlation coefficients were calculated. The mean inter-item correlations (MIC) were also calculated. These analyses were performed using SPSS (SPSS Inc., Chicago) and STATA (Stata Corp. LP, College Station). All other multivariate analyses were performed using Mplus (Muthén & Muthén, 1998).

Confirmatory factor analysis (CFA) was carried out using the weighted least squares means and variances adjusted (WLSMV) estimation, which is recommended for binary variables (Muthén & Muthén, 2001). Model fits for a one- and a two-factor (abuse and dependence) solution were examined using the comparative fit index (CFI), the root mean square error of approximation (RMSEA), the weighted root mean square residual (WRMR) and the Tucker-Lewis index (TLI). Recommended cut-off points for these measures are as follows: CFI ≥ 0.96, RMSEA ≤ 0.05, WRMR ≤ 0.9 and TLI ≥ 0.95 (Yu, 2002).

Latent class analysis (LCA) was applied to test whether groups of individuals with similar response patterns exist. The optimal number of classes was determined based on the Bayesian information criterion (BIC) (Nylund, Asparouhov, & Muthén, 2007). Analyses were performed until a relative minimum of the BIC value was identified. For graphical presentation, the estimated criterion probabilities for the seven dependence criteria and the four abuse criteria were merged. Expectation values representing the expected sum score of abuse ($E_A$) and dependence ($E_D$) criteria in each class were based on the number of criteria (four for abuse; seven for dependence) multiplied by the mean estimated probability per class.

Whereas LCA and CFA identify groups of individuals (classes) and groups of items (factors), latent class factor analysis (LCFA) is able to combine both approaches. Hence, LCFA can help identify response patterns under the assumption of a certain factor structure. Analyses were performed for an increasing number of latent classes combined with one or two factors (abuse and dependence). Model selection was based on the lowest BIC. LCFA was performed according to Clark et al. (2009), using model FMM-1. Covariate effects (metric age, sex, ordinal education) were modelled as regression paths on the latent classes. No cross-comparison with LCA and CFA is admissible, since no covariates were included within CFA and LCA.

## Results

Table 1 illustrates the criterion endorsement separately for abuse and dependence and for the URS and the RS. As expected, the prevalence rates were higher in the RS (1.7%–44.6%) than in the URS (0.4%–7.0%). Table 2 shows a cross-tabulation of abuse and dependence sum scores indicating a zero class of 6,104 individuals.

**Table 1**

*Criterion endorsement rates by sample*

|  | URS (n = 7,804) | RS (n = 1,700) |
|---|---|---|
| Neglect roles (A1) | 0.7 | 3.7 |
| Hazardous use (A2) | 2.0 | 7.8 |
| Legal problems (A3) | 0.4 | 1.7 |
| Social/interpersonal problems (A4) | 3.8 | 14.8 |
| Tolerance (D1) | 6.2 | 39.2 |
| Withdrawal (D2) | 0.6 | 3.7 |
| Larger/longer (D3) | 6.7 | 42.8 |
| Quit/control (D4) | 3.2 | 20.5 |
| Time spent (D5) | 7.0 | 44.6 |
| Impaired activities (D6) | 0.8 | 4.9 |
| Use despite problems (D7) | 1.7 | 11.4 |

URS unrestricted sample. RS restricted sample.

**Table 2**

*Cross-tabulation of abuse and dependence criteria sum scores*

| Abuse criteria | Dependence criteria | | | | | | | | Total |
|---|---|---|---|---|---|---|---|---|---|
|  | ∑=0 | ∑=1 | ∑=2 | ∑=3 | ∑=4 | ∑=5 | ∑=6 | ∑=7 |  |
| ∑=0 | 6,104 | 602 | 196 | 63 | 23 | 7 | 2 | 0 | 6,997 |
| ∑=1 | 117 | 76 | 43 | 38 | 12 | 7 | 1 | 1 | 295 |
| ∑=2 | 22 | 7 | 15 | 9 | 10 | 3 | 3 | 1 | 70 |
| ∑=3 | 3 | 1 | 1 | 3 | 1 | 1 | 2 | 1 | 13 |
| ∑=4 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 4 |
| Total | 6,247 | 686 | 255 | 114 | 46 | 18 | 9 | 4 | 7,379 |

Total number of cases (n = 7,379) deviates from the analytical sample (n = 7,804) due to the fact that sum scores for abuse and dependence can only be calculated for individuals with complete response on all criteria. The number of individuals without complete response is 425. The zero-class is represented by cell (0,0) = 6,104 individuals.

Table 3 shows the tetrachoric correlation matrix. While in the URS all correlations were significant, in the RS only 29 out of 55 correlations were significant ($p < .05$) and seven inter-criterion correlations turned negative. The mean inter-item correlation was higher in the URS (MIC = .57) than in the RS (MIC = .21). In particular, the correlation among dependence criteria and across scales decreased when the zero class was excluded. The mean correlation among abuse criteria was not as much affected by the selection of the sample.

Confirmatory factor analyses (CFA) in both samples showed higher model fit for the two-factor solution than for the one-factor solution. However, the difference in the URS was small, suggesting a one-factor solution due to parsimony (Figure 1). Factor loadings and model fit were consistently higher in the URS than in the RS. Based on a two-factor solution, the correlation between the factors resulted in high obliqueness in the URS ($r = .85$) and near-orthogonality in the RS ($r = .13$). No residual correlations were specified within the models.

**Table 3**

*Tetrachoric correlation matrix of abuse and dependence criteria by sample*

|  | A1 | A2 | A3 | A4 | D1 | D2 | D3 | D4 | D5 | D6 | D7 |
|---|---|---|---|---|---|---|---|---|---|---|---|
|  |  |  |  |  |  | ↙ Restricted sample (*n* = 1,700) |  |  |  |  |  |
| Neglect roles (**A1**) |  | .37 | .45 | .52 | .14 | .45 | .25 | .25 | .06 | .54 | .54 |
| Hazardous use (**A2**) | .62 |  | .48 | .18 | -.16 | .11 | .04 | -.11 | .01 | .31 | .13 |
| Legal problems (**A3**) | .62 | .67 |  | .58 | -.03 | .19 | .06 | .09 | .13 | .44 | .47 |
| Social/interpersonal problems (**A4**) | .74 | .56 | .76 |  | -.11 | .31 | .04 | .13 | -.17 | .28 | .33 |
| Tolerance (**D1**) | .53 | .38 | .37 | .49 |  | .01 | .00 | -.16 | -.15 | .08 | .18 |
| Withdrawal (**D2**) | .64 | .41 | .41 | .60 | .43 |  | .35 | .30 | .09 | .55 | .62 |
| Larger/longer (**D3**) | .61 | .54 | .46 | .59 | .64 | .66 |  | .21 | .09 | .25 | .28 |
| Quit/control (**D4**) | .56 | .34 | .40 | .56 | .44 | .58 | .67 |  | -.15 | .29 | .22 |
| Time spent (**D5**) | .49 | .52 | .51 | .47 | .57 | .50 | .70 | .47 |  | .31 | .11 |
| Impaired activities (**D6**) | .71 | .57 | .62 | .59 | .49 | .71 | .61 | .58 | .66 |  | .62 |
| Use despite problems (**D7**) | .73 | .48 | .66 | .65 | .61 | .77 | .67 | .57 | .58 | .78 |  |
|  |  |  | ↗ Unrestricted sample (*n* = 7,804) |  |  |  |  |  |  |  |  |

Significant correlations ($p < .05$) in bold-face.

**Figure 1**

*Results of confirmatory factor analyses (CFA) for two- and one-factor solution by sample*
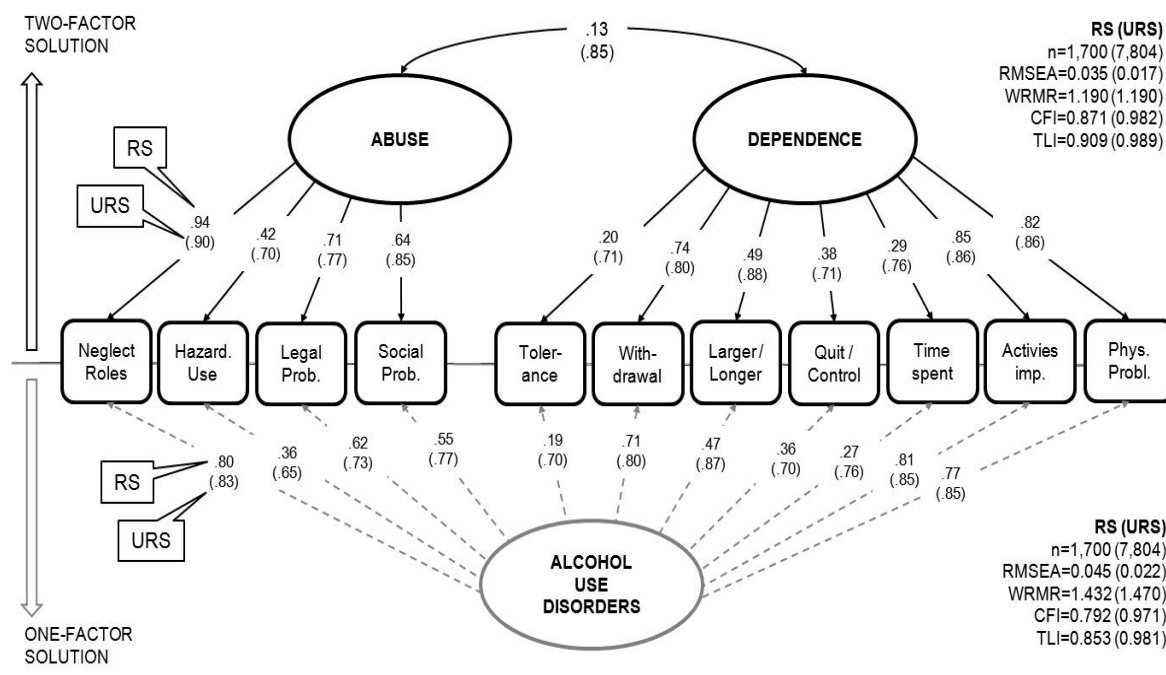


Standardized coefficients.
RS restricted sample. URS unrestricted sample. RMSEA root mean square error of approximation. WRMR weighted root mean square residual. CFI comparative fit index. TLI Tucker Lewis Index.

Latent class analysis favored three-class solutions in both samples (Table 4). In the URS, all classes showed higher criterion endorsement probabilities and expectancy values for dependence than for abuse criteria, indicating quantitative differences in severity across classes (Figure 2). In the RS, however, the pattern of class 2 diverged from the other classes, representing qualitative differences. Table 5 illustrates the transition of cases according to their class membership in the URS and the RS. The majority of cases in class 3 in the URS were excluded in the RS because of no criterion endorsement (99.2%). A separate "abuse class" only emerged in the RS and was composed of individuals categorized as the "nearly zero class" within the URS (96.8%). A distinct "dependence class" was visible only within the RS and was formed by individuals categorized as class 1 in the URS (88.5%).

LCFA including covariates (age, sex, education) favored a two-factor and five-class solution over all other combinations of one- or two-factor solutions with any amount of classes (Table 4; bottom). The best solution is depicted in Figure 3, in which classes are ordered according to their size. Class membership was highly related to covariates. The "dependent" (3), "premorbid dependent" (2) and "senior dependent" (5) classes all had below average education according to the standardized deviation from the average (z). The "zero class" (1) and the "abuse class" (4) had average and above average education levels. The only class with a majority of females was the "zero class" (59%). The "abuse class" was older (average age 47.9 years) than the "dependence class" (average age 31.9 years).

**Table 4**

*Model fit of Latent Class Analyses (LCA) and Latent Class Factor Analysis (LCFA)*

| | URS | | RS | |
|---|---|---|---|---|
| **Latent Class Analysis (LCA)[1]** | **BIC** | **Entropy** | **BIC** | **Entropy** |
| 1-class solution | 12,764 | 100.0 | 21,294 | 100.0 |
| 2-class solution | 12,283 | 89.0 | 18,090 | 90.5 |
| 3-class solution | 12,272* | 81.7 | 17,953* | 88.2 |
| 4-class solution | 12,297 | 80.5 | 17,970 | 90.1 |

| | One factor | Two Factors | |
|---|---|---|---|
| **Latent Class Factor Analysis (LCFA)[2]** | **BIC** | **BIC** | |
| 2 classes | 16,677 | 16,677 | |
| 3 classes | 16,386 | 16,355† | |
| 4 classes | 16,386 | 16,409 | |
| 5 classes | 16,431 | 16,303* | |
| 6 classes | 16,423 | 16,366 | |
| 7 classes | 16,445 | 16,420 | |

URS unrestricted sample. RS restricted sample. BIC Bayesian information criterion.
*best model according to BIC.
[1] LCA without covariates. [2] LCFA including covariates (age, sex, SES) for URS only.
[1,2] BICs of LCA and LCFA cannot be compared
† local minimum

**Table 5**

*Cross-tabulation of latent class membership in the unrestricted and restricted sample*

| | RS | | | | |
|---|---|---|---|---|---|
| **URS** | **Class 1 ("dependence")[a]** | **Class 2 ("abuse")[a]** | **Class 3 ("early dependence")[a]** | **Excluded[a]** | **Total[a]** |
| Class 1 ("abuse and dependence")[a] | **69** **(88.5%)** | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 69 (0.9%) |
| Class 2 ("abuse and dependence")[a] | 9 (11.5%) | 6 (3.2%) | 591 (58.5%) | 42 (0.6%) | 648 (8.3%) |
| Class 3 ("nearly zero")[a] | 0 (0.0%) | **180** **(96.8%)** | 419 (41.4%) | **6,488** **(99.4%)** | 7,087 (90.8%) |
| Total[β] | 78 (1.0%) | 186 (2.4%) | 1,010 (12.9%) | 6,530 (83.7%) | 7,804 (100.0%) |

[a] column percentages. [β] row percentages.
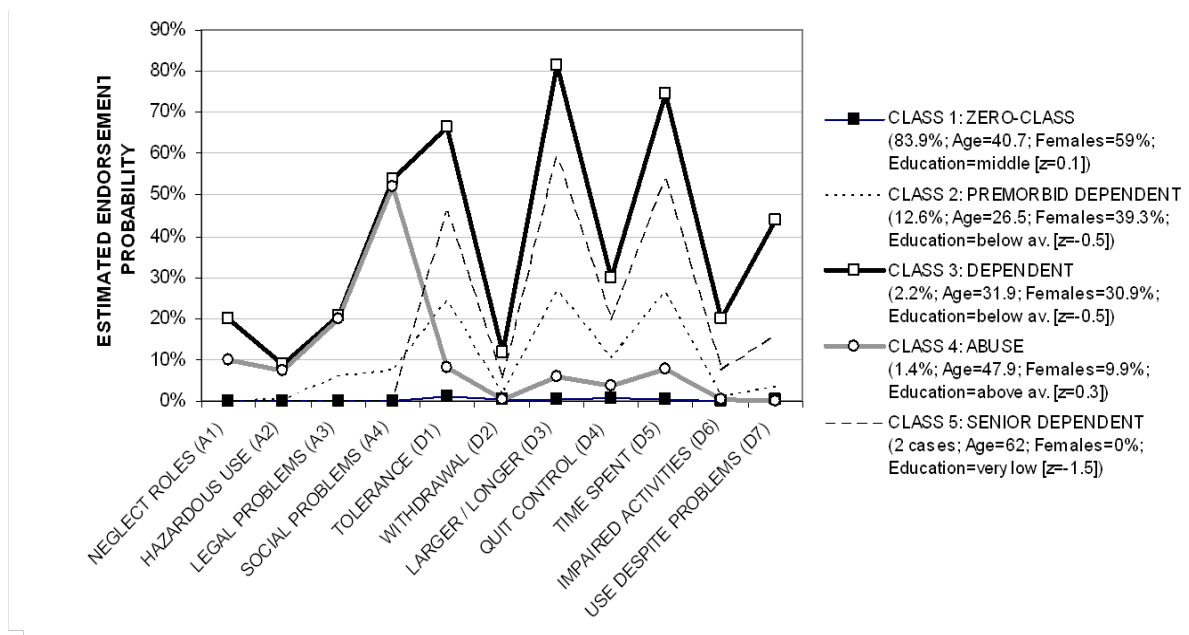URS unrestricted sample. RS restricted sample.

**Figure 2**

*Results of latent class analyses (LCA) by sample: Mean estimated criterion probabilities for dependence and abuse*
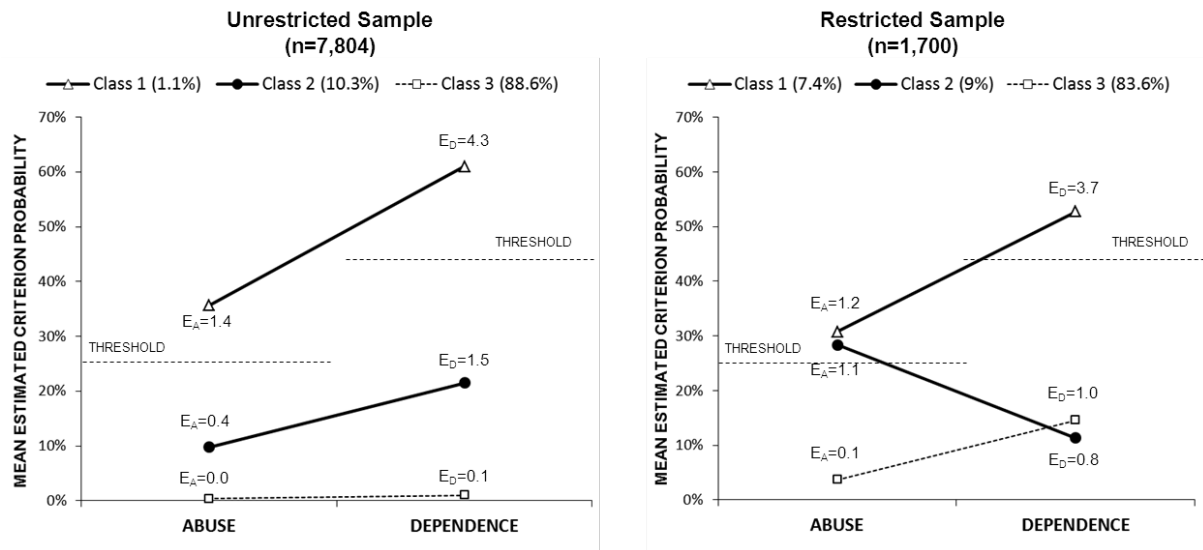


EA Expectation value for abuse criteria;  ED Expectation value for dependence criteria.  Diagnosis thresholds for abuse (1 criterion; *p*=25%) and dependence (3 criteria; *p*=43%) are indicated by dashed lines.

**Figure 3**

*Results of Latent Class Factor Analysis (LCFA)*



## Discussion

In an unrestricted German general population sample, the prevalence of DSM-IV AUD criteria was low, and the majority of individuals did not endorse any criterion. Different dimensionality analyses were affected by the inclusion of this large zero class. Restricting the analysis by including only individuals endorsing at least one criterion, or by applying hybrid models (identifying the zero class), led to findings that supported a structure of the diagnostic criteria with at least two dimensions.

When individuals without criterion endorsement were excluded, factor analyses (CFA) resulted in higher heterogeneity (more factors).  The reduction of homogeneity in the RS was related to a reduction of inter-item correlations.  The correlation matrix revealed that abuse criteria had more positive correlations among each other than with dependence criteria.  It can be assumed that the inclusion of individuals with no endorsed criteria in a sample "deteriorates" the covariance matrix and, therefore, the factor structure.  This may be the reason why previous studies using general population samples did not find abuse and dependence as separable factors

and preferred a more parsimonious one-factor solution (Hasin & Beseler, 2009; Keyes et al., 2011; Mewton, Slade, McBride, Grove, & Teesson, 2011; Proudfoot et al., 2006; Saha et al., 2006; Slade et al., 2009). If analyses had been applied to the URS only, the present study would have also accepted one-dimensionality due to parsimony.

The CFA also demonstrated that the correlation between abuse and dependence in a two-factor solution is only high in the URS. This high correlation had been responsible for the acceptance of a one-factor solution in previous studies (Keyes et al., 2011; Proudfoot et al., 2006). In the RS, however, a nearly orthogonal (zero) correlation was found, making a one-factor solution due to parsimony unacceptable. The CFA in the RS showed low factor loadings of the criteria "tolerance," "time spent" and "hazardous use." This result indicates that further factors other than abuse and dependence could emerge. Particularly, the finding on "tolerance" is in line with previous studies, such as Langenbucher et al. (2004), who dropped the criterion from the analysis due to a poor fit in the one-factor model. Similarly, Grant et al. (2007) found "tolerance" as a third separate factor in addition to abuse and dependence. Earlier problems with the criterion "legal problems" could not be replicated in the current study (Saha et al., 2006).

Latent class analysis resulted in different patterns, depending on the sample. Based on the URS, the parallel shift of classes with respect to criterion endorsement probabilities indicates that classes differ in terms of quantity (i.e., severity), rather than quality (i.e., patterns). This is in line with previous LCA studies on alcohol (Bucholz et al., 1996; Kuo, Aggen, Prescott, Kendler, & Neale, 2008; Muthén, 2006) and other substance use disorders (Agrawal, Lynskey, Madden, Bucholz, & Heath, 2007; Chung & Martin, 2005; Grant et al., 2006). Based on the RS, however, a qualitative interaction between abuse, dependence and class membership emerged. Hence, a separate "abuse" class could only be identified in the RS. It may be interpreted that the presence of a predominant zero class can distort the distinction of subtle class differences. Particularly, classes related to few criteria and low endorsement, such as abuse, seem to be prone to absorption by the "noise" of a large zero class.

The LCFA corroborates the results of LCA and CFA. The analysis identified not only two distinguishable factors, but also different response patterns according to abuse and dependence. The results contradict the assumption of abuse being a prodromal phase of dependence, since the "abuse class" (class 4) is older and shows a different pattern of covariates than the "dependence classes" (classes 2, 3, 5). This is in line with findings of earlier prospective studies (Grant, Stinson, & Harford, 2001; Hasin, Grant, & Endicott, 1990; Hasin, Van Rossem, McCloud, & Endicott, 1997; Schuckit et al., 2001; Schuckit et al., 2008). The fact that the abuse class is not parallel with the other classes in the LCFA, but intersects with them, throws into question the graded severity assumed by the revised diagnoses in DSM-5. Only one class ("premorbid dependent"; class 2) may correspond to a group remaining below the threshold of both abuse and dependence (i.e., "diagnostic orphans") (Mewton et al., 2011; Pollock & Martin, 1999; Ray, Miranda, Chelminski, Young, & Zimmerman, 2008). However, this group has low endorsement on the abuse

items; rather, its covariate profile is similar to a prodromal phase of dependence, apart from age (individuals in class 2 were younger). The small class 5 ("senior dependent") may be considered an artefact, but results in a better model fit than four- and six-class solutions. There is no *a priori* knowledge about the reliability of such small classes.

Criterion endorsement rates in our study were 0.6% to 7.0% in the URS and 1.7% to 44.6% in the RS. Compared to other studies, our URS was similar to general population samples used by Saha et al. (2006) (1.1%-12.9%), Saha, Stinson, and Grant (2007) (1.2%-13.8%), and Proudfoot et al. (2006) (1.1%-15.0%). All of these studies confirmed a one-dimensional structure of AUD criteria. The endorsement rates in the RS were comparable to Kahler and Strong (2006) (0.6%-30.5%), Hasin and Beseler (2009) (5.1%-39.4%), and Keyes et al. (2011) (1.6%-25.5%). Interestingly, these studies also confirmed a one-factor solution. This might be due partly to differences in the applied methods (e.g., "discontinuity approach") and the acceptance of one-dimensionality due to parsimony. Some recent studies have used clinical samples in order to get a more restricted population. However, these samples were not specific for alcohol-related disorders and also included other substance use disorders. These clinical samples therefore also had very low criterion endorsement rates—for example, Borges et al. (2010; 6.6% to 16.0%) or Hasin, Fenton, Beseler, Park, and Wall (2012; zero class for alcohol about 42%). Compared to Nelson et al. (1999), who tested the effect of sample composition on dimensionality, our RS was comparable to the full sample of current alcohol users in this earlier work (criterion endorsement 13.5% to 50.4%). Though we identified more than one factor in our RS, Nelson and colleagues found two factors only when restricting their sample even more (criterion endorsement rates for this "trimmed" sample were not reported).

The inclusion of covariates in the LCFA, and their exclusion in LCA and CFA, needs further explanation. The exclusion is justified by the fact that we wanted to produce results that were comparable with those of earlier studies, most of which did not include covariates (Hasin & Beseler, 2009; Keyes et al., 2011; Proudfoot et al., 2006; Saha et al., 2006; Slade et al., 2009). As well, the inclusion of covariates in CFA is equivalent to a MIMIC approach (Hardford & Muthén, 2001) and these models tend to result in more-dimensional structures (Grant et al., 2007; Hardford & Muthén, 2001; Simons et al., 2009). Consequently, the inclusion of covariates in the CFA and LCA led to a more-dimensional model even in the URS (results not shown), undermining the effect of the zero class. On the other hand, differences in covariates between URS and RS might be responsible for different factor analytic results. This led us to include the covariates in the final LCFA model based on the URS, which resulted in the identification of the zero class and abuse and dependence as distinct classes. Due to the different approaches, however, fit indices of the models cannot be compared.

Limitations of this study mainly concern the generalizability of the empirical findings. (1) It cannot be excluded that the results of the dimensionality analyses are due to the features of the German sample. Data from previous studies may be reanalysed excluding individuals without criterion endorsement. (2) It may be argued that dimensionality analyses should be conducted in the entire

population, because each individual should be assigned a meaningful score. However, according to Herrmann (1969) and Guilford (1959), a dimension may not necessarily be general or universal for the total population. More specifically, Cattell (1965) argues that it is not appropriate to test the dimensionality of highly specific characteristics in samples where the majority of members are not affected by the property in question. (3) Even in the RS there are still criteria with very low prevalence, including "legal problems" (1.7%) or "neglect roles" (3.7%). In item analysis these items would usually be excluded. Although only individuals with at least one endorsed criterion were included in the RS, these items may still inflate homogeneity. (4) Selection bias may have occurred; specifically, heavy drinkers with high probability of reporting AUD criteria may be underrepresented.

Our analyses confirm the assumption that a large zero class affects the dimensionality of AUD criteria, resulting in an increased homogeneity and, consequently, fewer dimensions. Limiting the analyses to an RS seemed not only to reduce homogeneity but to produce qualitatively different response profiles. With regard to the revision of the diagnostic system, our findings do not support the combination of abuse and dependence into a single diagnostic category of alcohol use disorder. Critique of the single category also came from two recent studies that found two distinct dimensions of alcohol-related problems assessed with the symptom items of the Alcohol Use Disorders Identification Test (AUDIT; Peng, Wilsnack, Kristjanson, Benson, & Wilsnack, 2012; Rist, Glockner-Rist, & Demmel, 2009). Results based on methods that consider person and item characteristics simultaneously also challenge the present DSM-5 AUD diagnostics, suggesting the necessity of either adding a further category or lowering the threshold for dependence.

# References

Agrawal, A., Lynskey, M. T., Madden, P. A., Bucholz, K. K., & Heath, A. C. (2007). A latent class analysis of illicit drug abuse/dependence: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Addiction, 102,* 94–104.

Baillie, A. J., & Teesson, M. (2010). Continuous, categorical and mixture models of DSM-IV alcohol and cannabis use disorders in the Australian community. *Addiction, 105,* 1246–1253.

Borges, G., Ye, Y., Bond, J., & Cherpitel, C. J., Cremonte, M., Moskalewicz, J., . . . Rubio-Stipec, M. (2010). The dimensionality of alcohol use disorders and alcohol consumption in a cross-national perspective. *Addiction, 105,* 240–254.

Bucholz, K., Heath, A. C., Reich, T., Hesselbrock, V. M., Kramer, J. R., Nurnberger, J. I. Jr., . . . Schuckit. M. A. (1996). Can we subtype alcoholism? A latent class analysis of data from relatives of alcoholics in a multicenter family study of alcoholism. *Alcoholism: Clinical and Experimental Research, 20,* 1462–1471.

Cattell, R. B. (1965). Factor theory psychology: A statistical approach to personality. In W. S. Sahakian (Ed.), *Psychology of personality: Readings in theory* (pp. 388–414). Chicago, IL, United States: Rand McNally.

Clark, S. L., Muthén, B., Kaprio, J., D'Onofrio, B. M., Viken, R., Rose, R. J., & Smalley, S. L. (2009). Models and strategies for factor mixture analysis: Two examples concerning the structure underlying psychological disorders. Retrieved from https://www.statmodel.com/download/FMA%20Paper_v142.pdf.

Chung, T., & Martin, C. S. (2005). Classification and short-term course of DSM-IV cannabis, hallucinogen, cocaine, and opioid disorders in treated adolescents. *Journal of Consulting and Clinical Psychology, 73,* 995–1004.

Edwards, G., & Gross, M. M. (1976). Alcohol dependence: Provisional description of a clinical syndrome. *British Medical Journal, 1,* 1058–1061.

Edwards, G. (2012). "The evil genius of the habit": DSM-5 seen in historical context. *Journal of Studies on Alcohol and Drugs, 73,* 699–701.

Gelhorn, H., Hartman, C. A., Sakai, J. T., Stallings, M. C., Young, S. E., & Rhee, S. H. (2008). Towards DSM-IV: An item response theory analysis of the diagnostic process for DSM-IV alcohol abuse and dependence in adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 47,* 1329–1339.

Grant, B. F., Hardford, T. C., Muthén, B. O., Hsiao-Ye, Y. E., Hasin, D. S., & Stinson, F. S. (2007). DSM-IV alcohol dependence and abuse: Further evidence of validity in the general population. *Drug and Alcohol Dependence, 86,* 154–166.

Grant, B. F., Stinson, F. S., & Harford, T. C. (2001). Age at onset of alcohol use and DSM-IV alcohol abuse and dependence: A 12-year follow-up. *Journal of Substance Abuse, 13,* 493–504.

Grant, J. D., Scherrer, J. F., Neuman, R. J., Todorov, A. A., Price, R. K., & Bucholz, K. K. (2006). A comparison of the latent class structure of cannabis problems among adult men and women who have used cannabis repeatedly. *Addiction, 101,* 1133–1142.

Guilford, J. P. (1959). *Personality*. New York, NY, United States: McGraw Hill.

Hardford, T. C., & Muthén, B. O. (2001). The dimensionality of alcohol abuse and dependence: A multivariate analysis of DSM-IV symptom items in the National Longitudinal Survey of Youth. *Journal of Studies on Alcohol, 62,* 150–157.

Hasin, D., Van Rossem, R., McCloud, S., & Endicott, J. (1997). Alcohol dependence and abuse diagnoses: Validity in community sample heavy drinkers. *Alcoholism: Clinical and Experimental Research, 21,* 213–219.

Hasin, D. S. & Beseler, C. L. (2009). Dimensionality of lifetime alcohol abuse, dependence and binge drinking. *Drug and Alcohol Dependence, 101,* 53–61.

Hasin, D. S., Fenton, M. C., Beseler, C., Park, J. Y., & Wall, M. M. (2012). Analyses related to the development of DSM-5 criteria for substance use related disorders: 2. Proposed DSM-5 criteria for alcohol, cannabis, cocaine and heroin disorders in 663 substance abuse patients. *Drug and Alcohol Dependence, 122,* 28–37.

Hasin, D. S., Grant, B., & Endicott, J. (1990). The natural history of alcohol abuse: Implications for definitions of alcohol use disorders. *American Journal of Psychiatry, 147,* 1537–1541.

Herrmann, T. (1969). *Lehrbuch der empirischen Persönlichkeitsforschung* [Textbook of empirical

personality research] (2nd ed.). Göttingen, Germany: Hogrefe.

Kahler, C. W., & Strong, D. R. (2006). A Rasch model analysis of DSM-IV alcohol abuse and dependence items in the National Epidemiological Survey on Alcohol and Related Conditions. *Alcoholism: Clinical and Experimental Research, 30,* 1165–1175.

Keyes, M., Krueger, R. F., Grant, B. F., & Hasin, D. S. (2011). Alcohol craving and the dimensionality of alcohol disorders. *Psychological Medicine, 41,* 629–640.

Kraus, L., & Baumeister, S. E. (2008). *Studiendesign und Methodik des Epidemiologischen Suchtsurveys 2006* [Study design and methodology of the 2006 Epidemiological Survey of Substance Abuse]. *Sucht, 54,* S6–S15.

Kuo, P. H., Aggen, S. H., Prescott, C. A., Kendler, K. S., & Neale, M. C. (2008). Using a factor mixture modeling approach in alcohol dependence in a general population sample. *Drug and Alcohol Dependence, 98,* 105–114.

Langenbucher, J. W., Labouvie, E., Martin, C. S., Sanjuan, P. M., Bavly, L., & Kirisci, L. (2004). An application of item response theory analysis of alcohol, cannabis, and cocaine criteria in DSM-IV. *Journal of Abnormal Psychology, 113,* 72–80.

Lynskey M. T, & Agrawal A. (2007). Psychometric properties of DSM assessment of illicit drug abuse and dependence: Results from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *Psychological Medicine, 37,* 1345–1355.

Martin, C. S., Chung, T., & Kirisci, L. (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.

Mewton, L., Slade, T., McBride, O., Grove, R., & Teesson, M. (2011). An evaluation of the proposed DSM-5 alcohol use disorder criteria using Australian national data. *Addiction, 106,* 941–950.

Muthén, B. (2006). Should substance use disorders be considered as categorical or dimensional? *Addiction, 101* (Suppl 1), 6–16.

Muthén, B. O. (1995). Factor analysis of alcohol abuse and dependence symptom items in the 1988 National Health Interview survey. *Addiction, 90,* 637–645.

Muthén, L. K., & Muthén, B. O. (1998). *Mplus Users' Guide*. Los Angeles, CA, United States: Author.

Muthén, L. K., & Muthén, B. O. (2001). *Mplus. Statistical analysis with latent variables. User's guide (v2.01).* Los Angeles, CA, United States: Author.

Nelson, C. B., Rehm, J., Üsün, T. B., Grant, B., & Chatterji, S. (1999). Factor structures for DSM-IV substance disorder criteria endorsed by alcohol, cannabis, cocaine and opiate users: Results from the WHO reliability and validity study. *Addiction, 94,* 843–855.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14,* 569.

Peng, C.-Z., Wilsnack, R. W., Kristjanson, A. F., Benson, P., & Wilsnack, S. C. (2012). Gender differences in the factor structure of the Alcohol Use Disorders Identification Test in multinational general population surveys. *Drug and Alcohol Dependence, 124,* 50–56.

Piontek, D., Kraus, L., Müller, S., & Pabst, A. (2010). To what extent do age, period, and cohort patterns account for time trends and social inequalities in smoking? *Sucht, 56,* 361–371.

Pollock, N. K. & Martin, C. S. (1999). Diagnostic orphans: Adolescents with alcohol symptoms who do not qualify for DSM-IV abuse or dependence diagnoses. *American Journal of Psychiatry, 156,* 897–901.

Proudfoot, H., Baillie, A. J., & Teeson, M. (2006). The structure of alcohol dependence in the community. *Drug and Alcohol Dependence, 81,* 21–26.

Ray, L. A., Miranda, R. Jr., Chelminski, I., Young, D., & Zimmerman, M. (2008). Diagnostic orphans for alcohol use disorders in a treatment-seeking psychiatric sample. *Drug and Alcohol Dependence, 96,* 187–191.

Regier, D. A., Kuhl, E. A., & Kupfer, D. J. (2013). The DSM-5: Classification and criteria changes. *World Psychiatry, 12,* 92–98.

Rist, F., Glockner-Rist, A., & Demmel, R. (2009). The Alcohol Use Disorders Identification Test revisited: Establishing its structure using nonlinear factor analysis and identifying subgroups of respondents using latent class factor analysis. *Drug and Alcohol Dependence, 100,* 71–82.

Saha, T. D., Chou, S. P., & Grant, B. F. (2006). Toward an alcohol use disorder continuum using item response theory: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Psychological Medicine, 36,* 931–41.

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.

Schuckit, M., Danko, G. P., Smith, T. L., Bierut, L., Bucholz, K., Edenberg, H. J. et al. (2008). The prognostic implications of DSM-IV abuse criteria in drinking adolescents. *Drug and Alcohol Dependence, 97,* 94–104.

Schuckit, M., Smith, T. L., Danko, G. P., Bucholz, K. K., Reich, T., & Bierut, L. (2001). Five-year clinical course associated with DSM-IV alcohol abuse or dependence in a large group of men and women. *American Journal of Psychiatry, 158,* 1084–1090.

Simons, J. S., Carey, K. B., & Wills, T. A. (2009). Alcohol abuse and dependence symptoms: A multidimensional model of common and specific etiology. *Psychology of Addictive Behaviors, 23,* 415–427.

Slade, T., Grove, R., & Teeson, M. (2009). A taxometric study of alcohol abuse and dependence in a general population sample: Evidence of dimensional latent structure and implications for DSM-V. *Addiction, 104,* 742–751.

Wittchen, H.-U. (1994). Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatric Research, 28,* 57–84.

Wittchen, H.-U., Beloch, E., Garczynski, E., Holly, A., Lachner, G., Perkonigg, A. . . . Zieglgänsberger, S. (1995). *Münchener Composite International Diagnostic Interview (M-CIDI), Paper-pencil 2.2, 2/95.* München: Max-Planck-Institut für Psychiatrie, Klinisches Institut.

Yu, C. Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes.* (Unpublished doctoral dissertation.) University of California, Los Angeles, CA, United States.