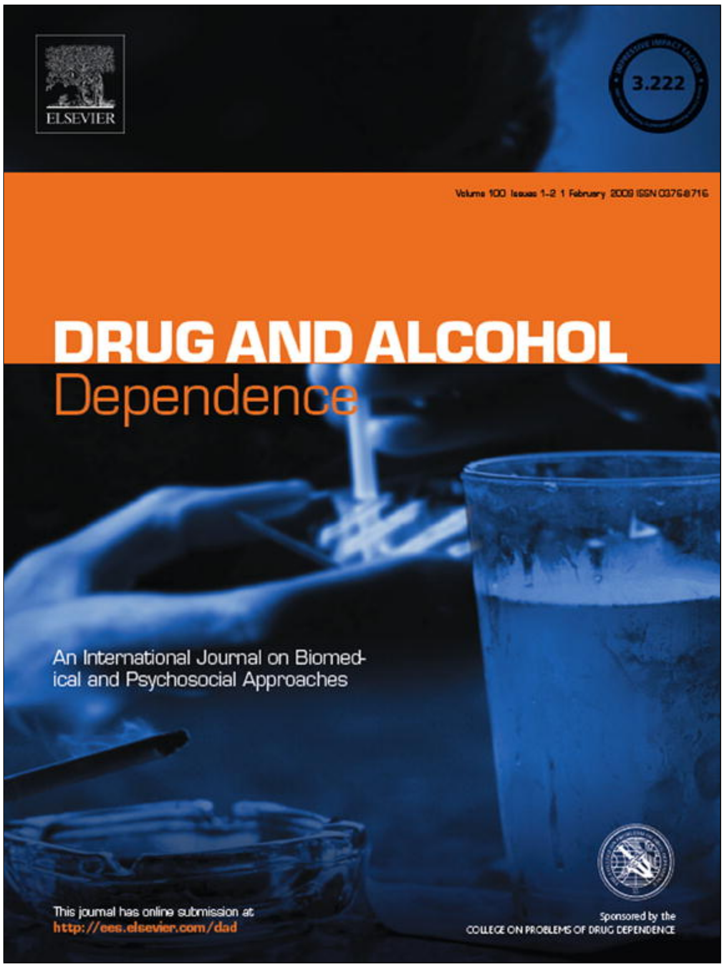
ELSEVIER
3.222
DRUG AND ALCOHOL
Dependence
An International Journal on Biomedical and Psychosocial Approaches
This journal has online submission at
http://ees.elsevier.com/dad
Sponsored by the
COLLEGE ON PROBLEMS OF DRUG DEPENDENCE

Drug and Alcohol Dependence 100 (2009) 71–82

Contents lists available at ScienceDirect

Drug and Alcohol Dependence

journal homepage: www.elsevier.com/locate/drugalcdep

The Alcohol Use Disorders Identification Test revisited: Establishing its structure using nonlinear factor analysis and identifying subgroups of respondents using latent class factor analysis

Fred Rist[a,*], Angelika Glöckner-Rist[b], Ralf Demmel[a]

[a] Department of Clinical Psychology, University of Münster, Fliednerstr. 21, 48149 Münster, Germany
[b] German Social Science Infrastructure Services, B1 15, 68072 Mannheim, Germany

ARTICLE INFO

Article history:
Received 20 March 2008
Received in revised form 1 September 2008
Accepted 9 September 2008
Available online 20 November 2008

Keywords:
Alcohol dependence
Factor structure
Gender differences
Hazardous alcohol use
Item response theory
Screening

ABSTRACT

Background: Previous research used principal components as well as exploratory and confirmatory factor analysis to establish continuous dimensions underlying answers to the 10-items of the Alcohol Use Disorders Identification Test (AUDIT). The majority of these studies conclude that one consumption dimension and an adverse consequences dimension explain the answers to the AUDIT sufficiently. However, most of the methods used presuppose normal answer distributions and linear relations between indicators and constructs, which are unrealistic assumptions for AUDIT answer.
Objectives: First, to investigate the continuous factor analytic structure underlying the answers to all AUDIT items. Second and third, to assess the impact of consumption as well as age and gender on AUDIT consequences dimension. Fourth and fifth, to categorize respondents into subgroups based on the AUDIT consequences items and adjusting the subgroups for differences in consumption, age and gender. Sixth, to describe the subgroups with respect to further adverse consequences of drinking.
Methods: Nonlinear factor and latent class factor analyses models were applied to the AUDIT answers of $N = 6259$ patients of 26 general practitioners in a city area in Germany. Consumption items as well as age and gender were included as predictors of answers to the AUDIT consequences items.
Results: Nonlinear factor analyses suggested two continuous correlated factors reflecting the adverse consequences of alcohol use: (1) harmful alcohol use, (2) alcohol dependence (aim 1). Consumption items did not prove to be reasonable construct indicators, but adverse consequences were predicted by consumption (aim 2), and also by age and gender (aim 3). Latent class factor analysis identified four subgroups based on the AUDIT consequences items (aim 4): one not affected (66%), and three subgroups defined by either harmful (15%) or dependent (9%), or combined harmful and dependent use (10%). These groups differed also with respect to further alcohol use consequences. Adjusting the subgroups for differences in consumption, age and gender (aim 5) reduced the non-affected subgroup and increased the subgroup with harmful and dependent use.
Conclusions: The AUDIT items cover three separable domains, i.e. consumption, harmful and dependent use, as originally intended. Hence, assessment of alcohol use does not substitute for assessing adverse consequences, as assumed in short versions of the AUDIT comprising only the AUDIT consumption items. Further, the dimensional as well as the LCFA subgroup solution imply that the respondents cannot be ordered along a single severity dimension without loss of information.

1. Introduction

The Alcohol Use Disorders Identification Test (AUDIT) was developed with the main objective to screen for alcohol use disorders (AUD) in various populations (Babor et al., 2001; Saunders et al., 1993). Therefore, seven of its items were selected from a larger pool of indicators assumed to reflect harmful and dependent drinking. Included were also three items assessing alcohol consumption behaviour in terms of quantity and frequency of drinking. Thus, the AUDIT can be employed to detect both AUD and at-risk alcohol consumption (ARC). This has been recognized as an asset in comparison to other screening instruments, which generally focus only on dependent and harmful drinking (Gordon, 2006).

* Corresponding author. Tel.: +49 251 83 34111; fax: +49 251 83 31331.
E-mail address: rist@psy.uni-muenster.de (F. Rist).

0376-8716/$ – see front matter © 2008 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.drugalcdep.2008.09.008

Initially, the AUDIT was meant to be combined with a clinical screening procedure including laboratory and physical examination evaluations. However, in most studies and in clinical routine, the AUDIT was used as a standalone screening procedure. For this purpose, the AUDIT has been psychometrically evaluated in numerous studies, applying methods and criteria of classical test theory (Reinert and Allen, 2007). In addition, several AUDIT answer sum scores have been proposed as cut scores to identify AUD and ARC. A sum score of 8 or more provides good sensitivity to detect AUD, but a cut score of 10 or more offers better specificity (Babor et al., 2001). Lower cut scores have been recommended for special populations or when screening for ARC is emphasized (Reinert and Allen, 2007). Studies of various (sub)populations (different national, ethnic, age, and gender groups) demonstrate convincingly that valid decisions regarding the risk of AUD and ARC can be based on AUDIT sum scores.

However, subsequent to a review of AUDIT research, Allen et al. (1997) identified construct validity and dimensionality of the AUDIT as questions in need of further enquiry. On the one hand, the constructors of the AUDIT intended to tap three different domains which should give rise to a corresponding dimensional structure of the AUDIT answers. On the other hand, a sum score would properly reflect the information contained in the AUDIT answers only if the AUDIT can be considered as unidimensional. One might feel tempted to take the often reported high Cronbach's Alpha values (> 0.80; Reinert and Allen, 2007) as indicating the high internal consistency of the AUDIT and as evidence for its unidimensionality. However, Cronbach's Alpha can be meaningfully computed only for homogeneous items (Cronbach, 1951; Cronbach and Shavelson, 2004). Consequently, unidimensionality cannot be derived from high alpha values: they do not preclude that the answers to the AUDIT items are determined by more than one latent trait.

Meanwhile, several studies have been conducted to clarify these issues by applying conventional linear dimensional analysis (Table 1). Two early studies applied Principal Component Analysis (PCA) to the data of alcohol dependent patients and concluded that one common component is sufficient to explain the variances of the AUDIT answers (El-Bassel et al., 1998; Skipsey et al., 1997). One component emerged also in a study of psychiatric patients in India (Carey et al., 2003). Maisto et al. (2000) submitted the data of primary care patients to PCA and linear Confirmatory Factor Analysis (CFA), after excluding item 10 because of its low factor loading (<0.40). The PCA suggested two components (3 consumption items; 7 consequences items), while the CFA led to solutions of comparable fit for both a two and a three factor structure (consumption, harmful, and dependent use). The authors accepted the two factor structure with the argument of parsimony. A similar two-dimensional solution was supported in seven more studies of different populations: mental health out-clinic patients (Karno et al., 2000), workers (Medina-Mora et al., 1998), undergraduates (O'Hare and Sherrer, 1999), both college students and substance use treatment outpatients (Shields et al., 2004), a Brazilian general population sample (Lima et al., 2005), and a Swedish general population study (Bergman and Källmen, 2002). Analyzing a population sample, Shevlin and Smith (2007) were the only ones to report a clearly better fit for a three- than for a two-dimensional solution. Nevertheless, they argued for two dimensions, as their two consequences factors harm and dependence did not correlate differently with various demographic and quality of health characteristics of their respondents.

To summarize the results of these studies: One factor or component was rarely found and appeared to be confined to populations with a higher prevalence of AUD (El-Bassel et al., 1998; Skipsey et al., 1997; Carey et al., 2003). Three dimensions were identified

Table 1
Studies applying principal component analyses or factor analyses to AUDIT answers.

Year of publication	Authors	Sample	Methods	Number of components/factors
1997	Skipsey et al.	82 patients (62% female)	PCA/Varimax	1 (64%)[a]
1998	Medina-Mora et al.	Male workers	PCA/Varimax	2
1998	El-Bassel et al.	Female drug dependents	PCA/Varimax	1 (67.3%)[a]
1999	O'Hare and Sheerer	312 undergraduates (36% female)	PCA/Varimax	2 (24% and 36%)[a]
2000	Karno et al.	197 mental health patients (14% female)	CFA 1 vs. 3 Fs PCA and EFA	3 better 1 2/2
2000	Maisto et al.	7035 primary care patients (47% female)	PCA/Varimax CFA ~ item 10	2 3 better 2
2001	Gmel et al.	10,321 health survey participants (51% female)	CFA 1 vs. 3 Fs CFA 4F, 3Fs for AUDIT 1–3 Analysis of tetrachoric correlations	3 better 1 4 better 3
2001	Kelly and Donovan	103 adolescents and young adults; (50% female)	CFA 1 vs. 2 vs. 3 Fs	2 + 3 better 1, but F2 + F3 of 3 model correlated 1.00
2002	Chung et al.	173 med patients (57% female)	CFA 1–3 Fs PCA/Varimax	2 better 1; no convergence for 3
2002	Bergman and Källmen	997 general population	EFA, CFA 1–3 Fs	3 comparable 2
2003	Carey et al.	1349 patients (30% female)		
2004	Shields et al.	465 students (59.1% female) 135 outpatients (26.9% female)	CFA 1–3 Fs	3 comparable 2
2005	Lima et al.	General population	EFA	2
2007	Doyle et al.	Alcohol-dependent patients 1337 from study COMBINE 1711 from study MATCH	CFA 1–3 Fs	3 comparable 2
2007	Shevlin et al.	7849 general population	CFA 1–3 Fs	3 better 2, but 2 retained

CFA: Confirmatory Factor Analysis; EFA: Exploratory Factor Analysis; PCA: Principal Factor Analysis.
[a] Explained variance.

repeatedly (Maisto et al., 2000; Bergman and Källmen, 2002; Shevlin and Smith, 2007). Since these solutions were either not statistically superior to two-dimensional ones, or the identified alcohol related consequences dimensions seemed redundant with respect to their relations to other health related variables (Shevlin and Smith, 2007), two-dimensional solutions were favored as the more parsimonious ones. However, Gmel et al. (2001) considered even four dimensions as necessary: Three of them were determinants of consumption, but only one was associated with harmful and dependent use (adverse consequences).

The question of the dimensionality of the AUDIT items appears thus to be settled: Obviously, one dimension captures the variance or covariance of the AUDIT answers insufficiently, while three or more dimensions corresponding to consumption, harmful and dependent use might overspecify them. Instead, AUDIT answers seem to be determined by two separable albeit correlated latent variables, consumption and adverse consequences. This view is supported by CFA of eight correlation matrices obtained in different studies, including also two new data sets from alcohol dependent patients according to the DSM-IV criteria from the COMBINE and the MATCH study (Doyle et al., 2007). These analyses again suggest a two factor solution, even for the AUDIT answers of samples with a high prevalence of AUD.

However, before this final verdict on the structure of the AUDIT is accepted, we should consider several limitations of the studies cited. First, most applied multivariate procedures which presuppose interval scaled variables, multi normally distributed answers, and linear relations between components or factors and item answers. The AUDIT items hardly fulfill these prerequisites: Most of them yield highly skewed distributions when answered in populations with low prevalence of AUD. Yet only Gmel et al. (2001) and Shevlin and Smith (2007) analyzed tetrachoric correlations. Thereby the skewness of AUDIT answer distributions can be taken into account, but not also possible nonlinear relations between AUDIT answers and their underlying AUDIT dimensions. Nonlinear relations are to be expected for the consumption items, however: Item 3 inquires about the frequency of binge drinking, defined as at least six standard drinks. It will be redundant to the combined item 1 (frequency) and item 2 (quantity), when quantity assessed by item 2 equals six or more standard drinks. Thus, the relation between these items will vary with the level of drinking. This numerical dependence among the consumption items will introduce nonlinear relations and will limit the precision of any conventional dimensional analysis. Also, it is questionable whether consumption items reflect the influence of an underlying consumption construct (Chin, 1998). Instead, they might be more reasonably considered as behaviour descriptors. Second, although more recent studies used CFA, which is definitely to be preferred to PCA when searching for underlying dimensions, they did not exploit an important advantage of these models over conventional exploratory linear models: Inspection and modeling of the residual structure, to detect additional substantial communalities between indicators not explained by factors explicitly defined by a CFA model.

Therefore, the first aim of this study was to determine the dimensional structure of the AUDIT answers applying nonlinear factor analysis (FA) models. This approach takes into account non-normal distributions of indicators and their nonlinear relations with dimensions. Application of these methods instead of linear analysis tools will avoid that fit indices and parameter estimations are biased as a consequence of erroneous normality and linearity assumptions (Vermunt and Magidson, 2004). This will yield more robust and trustworthy results than the analyses performed in previous AUDIT studies.

A closely related second aim was to clarify the status of the AUDIT consumption items in a continuous factor model, since it is dubious if they can reasonably serve as construct indicators in a FA model. Alternatively, they will be included as predictors of consequences items into a nonlinear FA model of the consequences.

Age and gender are ubiquitously related to alcohol use, ARC and AUD (Day and Homish, 2002). Therefore, the third aim of the study will be to investigate their influence on AUDIT consumption and consequences items. To this end, age and gender will be allowed to simultaneously predict consumption and consequences items in the nonlinear FA model of the consequences.

One class of such nonlinear FA models is the two item parameter (item discrimination and item difficulty) Item Response Theory (IRT) model. It was developed for a more appropriate handling of binary and polytomous items than possible with conventional linear psychometric methods of classical test theory. With the multidimensional normal ogive variant of this model (Bock et al., 1988; McDonald, 1996), Takane and DeLeeuw (1987) have established the equivalence of the common linear factor model and IRT models. Accordingly, the latter model also provides a factor score for each person on each dimension, and a discrimination or loading for each item on each dimension. In addition, for each item threshold values are computed which are related to the proportion of positive responses which, in classical test theory, is commonly referred to as the item difficulty. In contrast to the conventional linear model this requires that not only the second, but also first distribution moments are involved in the model, and that tetrachoric or polychoric instead of Pearson correlations are analyzed. (cf. Glöckner-Rist and Hoijtink, 2003). Parallel to the normal ogive models the logistic IRT models (Mislevy, 1986; Reckase, 1996) have been developed. They determine the conditional probability by integrating over the logistic instead of the normal distribution. Yet they lead to the same results with the only exception that their discrimination parameters are always about 1.7 times larger than the corresponding parameters of the normal ogive models. Thus, also the logistic IRT models can be reparameterized as factor models. Both model variants have been integrated in a generalized structural equation modeling (SEM) framework by Muthén (2002). They can be computed with the program Mplus (for more details, visit http://www.statmodel.com) relying on robust weighted least squares (normal ogive model) and maximum likelihood (logistic model) estimation. This allows exploiting all modeling advantages for nonlinear models which the SEM approach provided only for linear FA models, i.e. confirmatory specification of dimensions and residuals, as well as comparisons of alternatively specified models.

The fourth aim of this study was to categorize AUDIT respondents into discrete latent, i.e. not directly observed subgroups on the basis of their answer profiles. When the AUDIT is used for screening relying on a cutoff score, respondents are either placed into a group at risk for ARC or AUD, or a group not at risk. The AUDIT manual (Babor et al., 2001) suggests even more subgroups, when recommending different interventions for respondents falling into different ranges of AUDIT scores. Thus, if assigning respondents to subgroups is the central decision to be based on the AUDIT, then clustering methods categorizing respondents directly into distinct groups should be applied. Again methods for non-normal indicators and nonlinear indicator–factor relationships should be preferred for AUDIT answer profiles. Latent class analysis (LCA) is particularly suited for this task (Hagenaars and McCutcheon, 2006).

LCA assumes only nominally distributed latent class dimensions and binary or polytomous observations. Newer hybrid models combine LCA with discretized versions of the continuous factors of the nonlinear two parameter IRT model (Heinen, 1996; Magidson and Vermunt, 2001). Thereby they allow examining whether subgroups can be ordered along two or more discrete latent class factors. If this is the case, the resulting subgroups are derived by cross-classification according to the levels of the identified class dimensions. Thus, the notion of one or more construct facets influencing the item responses is retained for the formation of distinct latent classes by this modeling approach.

The fifth aim is to delineate the influence of consumption, age and gender on the formation of subgroups and to adjust subgroups for these background variables. This interest is based on the same reasoning as the inclusion of these covariates into the nonlinear continuous FA model.

The sixth and last aim is to illustrate differences between the obtained subgroups with reference to further adverse consequences of drinking, not used for the formation of the subgroups. Selected consequences from the Drinker Inventory of Consequences (DrInC; Miller et al., 1995) will be used to further characterize the subgroups.

2. Methods

2.1. Recruitment procedure

The data were collected as part of a randomized clinical trial to assess the efficacy of screening and brief intervention (SBI) for hazardous alcohol use conducted in general physicians' offices in Germany (Demmel et al., 2003). $N = 26$ general practitioners (GPs; 14 male, 12 female; mean age 46.4 years), located in offices in and close to the city of Essen, participated in the investigation. The GPs were affiliated to the medical faculty of the University of Duisburg-Essen by a teaching network. They were introduced to SBI in two workshops, and project research assistants conducted site visits to assure proper implementation of SBI in the offices. The recruitment period was terminated when no more new patients could be recruited and did not exceed one year. On each day during the recruitment period, the nurses in the participating offices handed the screening questionnaire to the first eight patients per day who had an appointment with the GP and had not filled out the questionnaire during a previous visit. The limit of eight patients should allow the physician to conduct the intervention with every patient who was assigned to receive an intervention following the screening procedure.

2.2. Participants

Inclusion criteria were: age between 18 and 60 years, native language German or at least three years of schooling in the German school system, and regular (no emergency) visit to a GP. Pregnant women were excluded. From 11,469 patients who were asked by the nurses to take part in the initial screening, 2820 (25%) patients refused. Of those 8649 patients who had answered the questionnaire, 6259 (72%) patients were finally included in the analysis sample. $N = 560$ (7%) had not signed the informed consent properly ($N = 120$), did not meet the inclusion criteria (mostly less than three years of German education, $N = 266$), or had not given information relevant for inclusion ($N = 174$). $N = 649$ (8%) had missing or inconsistent responses on the AUDIT, $N = 368$ (4%) on the DrInC questionnaire. Inconsistent answering was coded in the case of contradicting responses to the three consumption items, e.g. if the frequency of binging occasions (item 3) surpassed the total number of drinking occasions (item 1). $N = 813$ (9%) did not drink alcoholic beverages and thus were not included in the analyses.

The mean age of the analysis sample was 35.5 years. 51% were female, 50% were single, 39% were married and living together with their partner, 3% were married, but not living together, 8% were divorced and 0.7% widowed. Educational level according to secondary school achievement/university degree was lower for 27%, intermediary for 53%, and high for 20% of the sample.

2.3. Instruments

2.3.1. Alcohol Use Disorder Identification Test (AUDIT)

The AUDIT was translated into German with the aim to replicate both the item content and the answer categories of the English version in everyday language. It passed several text revisions and exploratory applications, but the original answer categories were retained. Using part of the study sample ($N = 2101$; all patients recruited between July and November 2002), sensitivity and specificity to detect ARC (defined as drinking above the limits of 20 g alcohol daily for women and 30 g daily for men) was determined. With optimal cut scores of 4 (women) and 7 (men), the sensitivity was 0.80 and the specificity 0.75 for both men and women (Rist et al., 2003).

2.3.2. Drinker Inventory of Consequences (DrInC)

Based on the psychometric characteristics of a German-language version (Laumeyer, 2003), five items of moderate difficulty were included in our screening questionnaire: (a) having driven under the influence of alcohol, (b) having done embarrassing things, (c) having experienced trouble sleeping after drinking, (d) having had hangovers, (e) having had to vomit after drinking.

2.4. Statistical analysis

2.4.1. Confirmatory Factor Analyses (CFA) to determine the continuous factorial structure of the AUDIT (aim 1)

All nonlinear CFA were computed with Mplus (Version 4.2) relying on the two parameter normal ogive IRT model and robust mean and variance (MV) adjusted weighted least square (WLSMV) estimation. Compared to the logistic variant of this model class, which is also provided by Mplus, the normal ogive model has the technical advantage that it does not require numerical integration and is therefore faster to compute.

To yield the dimensional solution which best explains the covariances of the AUDIT answers, models with one, two, and more dimensions were compared.

2.4.2. Entering consumption items as well as age and gender as covariates into a nonlinear FA model of the consequences (aim 2 and aim 3)

In the final model, consumption items were included either as construct indicators or merely as behaviour descriptors to predict the consequences. Subsequently, age and gender were included as predictors of both consumption and consequences. For this purpose, age was transformed into five intervals (18–24, 25–29, 30–39, 40–49, 50–60 years). For each covariate, we tested, whether it affected the answers to the consequences items (a) directly (direct effects), or (b) indirectly via the underlying dimension(s) specified in the CFA model (indirect effects).

For a statistical evaluation of model fit, χ^2-values for overall goodness-of-fit will be reported, although they too often suggest model rejections in large samples (Bollen, 1989). Model comparisons were conducted with a special χ^2-difference test for WLSMV provided by Mplus.

As descriptive fit indices, the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI), and the Root Mean Square Error of Approximation (RMSEA) are considered relying on common decision criteria. According to these, an acceptable and good fit, respectively, requires for the first two indices values >0.90 and >0.95, respectively, and for the RMSEA values <0.10 and close to 0.05, respectively (Bollen, 1989).

2.4.3. Latent Class Analyses (LCA) and Latent Class Factor Analyses (LCFA) to identify subgroups with distinct AUDIT consequences profiles (aim 4)

Only the consequences items were used as class indicators. Following a widely used strategy to determine the number of classes required explaining answer profile differences sufficiently, conventional LC models with an increasing number of nominal latent classes were compared. Subsequently, LCFA were performed to explore whether the subgroups identified in the first step are ordered along discretized class dimensions. Again, consumption items, gender and age were included as covariates in the finally accepted class solution. In addition, the resulting subgroups were compared with respect to the prevalence of further alcohol related consequences according to the DrInC questionnaire.

The LCA and the LCFA were conducted with Latent Gold (LG; Version 4.1), relying on maximum likelihood estimation. While these analyses could have been similarly performed with Mplus, LG offers the advantage of a menu-driven model specification, a convenient graphical output, and it provides linearly approximated standardized discrimination parameters, which can be interpreted like the familiar standardized factor loadings from the conventional linear factor model. Model comparisons were based on the Bayes information criterion (BIC). When different models are applied to the same data, lower levels of the BIC indicate better solutions. The adequacy of class solutions was further evaluated by considering a classification error (CE; Vermunt and Magidson, 2004). It summarizes how precisely respondents could be assigned to subgroups on average, given their conditional, i.e. class specific answer profiles. The closer this value is to 0, the better the classification. In addition, residual structures were investigated in order to uncover additional systematic influences on answer processes which had not been explicitly taken into account by the models.

2.4.4. Assessing the influence of consumption as well as age and gender on the formation of subgroups (aim 5)

Analogous to the continuous factor model, the class probabilities will be regressed on consumption items as well as age and gender (covariates). The associations between age, gender and consumption will be determined already by the continuous the covariates and the consequences will be similar to those obtained in the FA model, they will be introduced not sequentially but simultaneously. Adjusted subgroup sizes will then be compared to the original subgroup sizes.

2.4.5. Comparing the subgroups with respect to further adverse drinking consequences (aim 6)

The endorsement probabilities for five consequences from the DrInC for each subgroup will be provided and compared between groups.

2.4.6. Accounting for influences of GP office on structural and class analyses

The data were collected in 26 GP offices, thus one might consider a multilevel approach to take into account a potential influence of the place of data collection. While this is an appropriate consideration for the effect of the brief interventions conducted by the GPs, there were no a priori reasons why the GPs office should have affected AUDIT answers systematically. The recruitment of patients and the

Table 2
Content of AUDIT items and frequency of AUDIT responses in the total sample ($N = 6259$).

Items	Responses				
	0	1	2	3	4
(1) Alcohol use, frequency[a]	–	0.30	0.42	0.21	0.07
(2) Alcohol use, quantity[b]	0.52	0.28	0.12	0.05	0.03
(3) Alcohol use, binge drinking[a]	0.28	0.45	0.20	0.07	0.01
(4) Loss of control[c]	0.85	0.10	0.03	0.01	0.01
(5) Role failure[c]	0.94	0.05	<0.005	<0.005	<0.005
(6) Morning drinking[c]	0.99	0.01	0.00	0.00	0.00
(7) Guilt feelings[c]	0.86	0.12	0.02	0.01	0.00
(8) Blackouts[c]	0.87	0.12	0.01	0.00	0.00
(9) Injuries[d]	0.89	–	0.09	–	0.02
(10) Concern[d]	0.91	–	0.05	–	0.04

[a] 0: never; 1: monthly or less; 2: 2–4 times/month; 3: 2–3 times/week; 4: ≥4 times/week.
[b] 0: 1–2 drinks; 1: 3–4 drinks; 2: 5–6 drinks; 3: 7–9 drinks; 4: ≥10 drinks.
[c] 0: never; 1: less than monthly; 2: monthly; 3: weekly; 4: daily or almost daily.
[d] 0: no; 2: yes, but not in the last year; 4: yes, during the last year.

answering of the questionnaires were conducted in a highly standardized fashion alike in all GP offices. In addition, involving place of data collection as a covariate in the finally accepted dimensional and class models did speak as expected against a significant influence of this variable on the AUDIT answers.

3. Results

3.1. Answer frequencies for the AUDIT items

The mean AUDIT score for the final sample ($N = 6259$) was 5.12 ($SD = 4.03$). 8.6% of the women and 30.6% of the men screened positive when the commonly used cut score of eight was applied. The answer frequencies for the 10 AUDIT items are presented in Table 2. The most frequently reported adverse consequence was feelings of guilt, reported by 14% of the respondents, followed by blackouts (13%) and injury of self or others (11%).

Morning drinking (item 6) was affirmed by less than 1% of the respondents (50 men and 6 women). To obtain answer frequencies of at least five percent per cell, this item had to be excluded from statistical analyses, and for several other items answer categories had to be combined. However, the distribution of the answers to item 6 across identified subgroups will be reported. The number of answer categories for the frequency and the binging items was reduced to four (Table 2). The answers to the adverse consequences items (items 4–10) were treated as dichotomous (present/not present). For item 1, four categories were retained, as respondents who stated that they "never" drunk alcohol were not included in any analysis.

3.2. Confirmatory factor analyses to determine the continuous factorial structure underlying AUDIT answers

3.2.1. Inclusion of all AUDIT items as construct indicators (aim 1). Common statistical and descriptive indices of overall fit for the CFA models tested are listed in Table 3, the factor-item associations for different factor solutions in Table 4. According to a χ^2 difference test, a model with two separate dimensions for consumption and adverse consequences items (MF2) explains the AUDIT answer covariances significantly better than a unidimensional model (MF1). The two-dimensional solution replicates previous results (e.g., Shields et al., 2004), yielding a consumption factor and an adverse consequences factor with substantial loadings. Both factors are highly correlated (0.82). A model with three dimensions (MF3), however, fits yet significantly better, achieving in contrast to the first models not only an acceptable, but also a good fit. It introduces a third dimension of dependent use, as it separates the two items "loss of control" (item 4) and "role failure" (item 5) from the other adverse consequences items, which comprise a factor of harmful consequences.

Inspection of the residuals shows, however, that even the three-dimensional model leaves substantial portions of the answer

Table 3
Overall fit indices for all continuous factor models based on 9 AUDIT items.

Models	χ^{2}[a]	df	CFI[b]	TLI[c]	RMSEA[d]
All items as construct indicators					
MF1: one dimension	834.9	23	0.93	0.95	0.08
MF2: two dimensions	545.2	22	0.96	0.97	0.06
MF3: three dimensions	514.6	21	0.96	0.96	0.06
MF4: four dimensions	212.7	18	0.98	0.99	0.04
Consumption items as covariates					
MF1i: one dimension, indirect effects	318.0	22	0.97	0.98	0.05
MF1d: one dimension, direct effects	101.6	12	0.99	0.99	0.04
MF2i: two dimensions, indirect effects	201.0	18	0.98	0.98	0.04
MF2d: two dimensions, direct effects	74.1	11	0.99	0.99	0.03
Consumption items, age, and gender as covariates					
MF2i: two dimensions, indirect effects	262.8	20	0.98	0.97	0.04
MF2d: two dimensions, direct effects	68.1	16	1.00	0.99	0.02

[a] All values significant.
[b] CFI: Comparative Fit Index.
[c] TLI: Tucker Lewis Index.
[d] RMSEA: Root Mean Square Error of Approximation.

Table 4
Factorial loadings for different dimensional models (MF1–MF3; all items treated as construct indicators).

	MF1	MF2		MF3		
	F1	F1	F2	F1	F2	F3
Frequency	0.50	0.51	–	0.51	–	–
Quantity	0.65	0.66	–	0.67	–	–
Binge	0.84	0.92	–	0.92	–	–
Loss of control	0.71	–	0.74	–	0.86	–
Role failure	0.63	–	0.66	–	0.73	–
Guilt	0.74	–	0.76	–	–	0.76
Black outs	0.79	–	0.82	–	–	0.82
Injuries	0.66	–	0.68	–	–	0.68
Concern	0.74	–	0.76	–	–	0.76
Factor correlations						
F1 with F2		0.82			0.68	
F2 with F3					0.84	
F1 with F3					0.83	

$N = 6259$.

covariances of the quantity and the binge consumption item (0.19) as well as the quantity and the frequency consumption item (−0.39) unexplained. Residual correlations are often treated only as deviations from the model specifications, and are included merely to enhance model fit. In contrast, we followed the qualified recommendation to consider correlations of residuals as indicating an influence of further, not yet accounted for factors which should to be included as such in a model (Bollen, 1989; Tomarken and Waller, 2005; DeShon, 1998). The respective residual correlations suggested nonlinear covariation of the consumption items: Although the answers to all three correlate positively with a consumption factor established by models MF2 and MF3 (Table 4), the residuals of quantity and frequency correlate negatively. Thus, a reciprocal relation between them emerges and remains unexplained when a positive association between all three consumption items is accounted for by a common factor underlying their answers. Therefore, in a further model a fourth factor, called “consumption habits”, was introduced in order to capture this additional reverse relationship between quantity and frequency (MF4). This model explains the data significantly and noticeably better than the three-dimensional model (Table 3). Fig. 1 shows that frequency of drinking is associated negatively (−0.56) and quantity of drinking positively (0.46) with the consumption habits factor, while a substantial association of all three consumption items with the common consumption factor still holds (≥0.56).

Apparently, the complex relations of consumption items to each other as well as to the other AUDIT items are adequately captured neither by a two nor by a three-dimensional model. Even the addition of a fourth dimension does not prune several other residual correlations involving the items 4–10 and also one of the three consumption items. Therefore, a different model building strategy was employed to investigate their relations to the other AUDIT consequences items.

3.2.2. Inclusion of consumption items as predictors instead of indicators (aim 2). As before, a one and a two-dimensional model was tested. Both models assumed only the consequences items as indicators either of only one dimension of adverse consequences (MAF1) or of two dimensions of harmful and of dependent use (MAF2). The three consumption items were included in both models as covariates, i.e. as predictors of adverse consequences. We explored, to which extent they influenced the answers to the indicators of one construct simultaneously via the respective latent

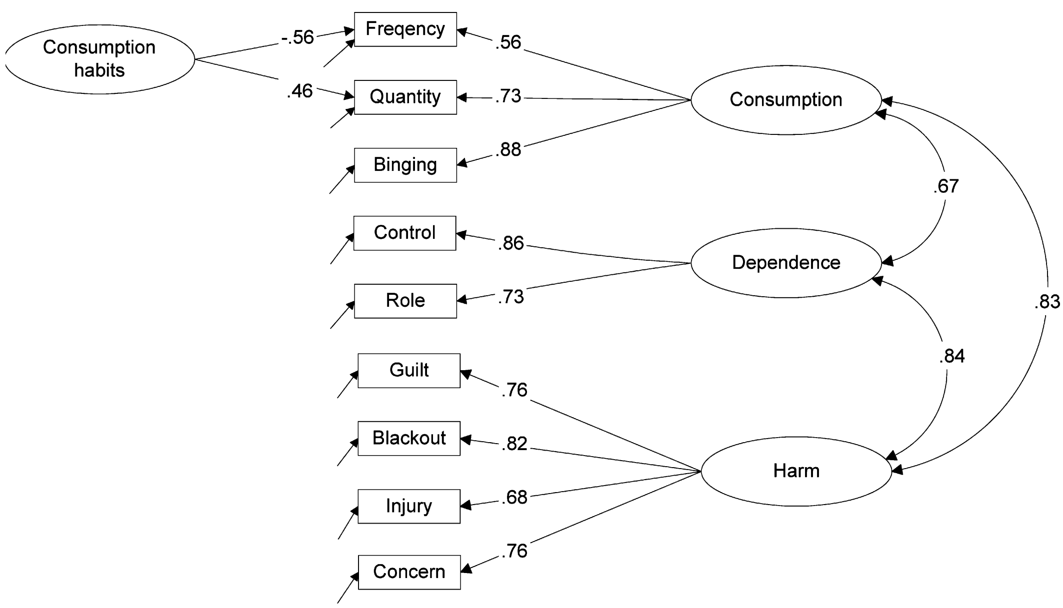

Fig. 1. Structural equation model MF4: Three dimensions assigned to consumption, dependence and harm items plus a fourth dimension to account for nonlinear relations among consumption items.

Table 5
Standardized regression coefficients for direct effects of AUDIT consumption items as well as age and gender on AUDIT adverse consequences.

Items	Covariates				
	Frequency	Quantity	Binging	Age	Gender[a]
Frequency	–	–	–	0.08	0.21
Quantity	0.07	–	–	−0.19	0.34
Binging	0.39	0.48	–	−0.17	0.37
Loss of control	–	–	0.56	–	–
Role failure	–	–	0.45	−0.09	–
Guilt feelings	0.06	–	0.28	−0.09	−0.13
Blackouts	–	0.07	0.37	−0.21	–
Injury	–	–	0.29	−0.24	–
Concern	0.07	–	0.23	–	–

[a] 0: female, 1: male.

variable (indirect effects), or directly (direct effects). According to the fit indices (Table 3), each model explained the item answer covariances better than even the four-dimensional model treating the consumption items as construct indicators. But according to a χ^2-difference test and the RMSEA, the model with a harmful and a dependent drinking factor again outperforms a model with only one common factor determining the answers to all adverse consequences items. Modeling the consumption items as behaviour descriptors thus improves model fit substantially, while preserving the two distinct facets of adverse consumption consequences (Table 4).

3.3. *Inclusion of age and gender as predictors of consumption and consequences (aim 3)*

According to significant standardized regression coefficients, being male (0.20), drinking more frequently (0.35), and drinking larger quantities (0.26) enhances the values of the respondents on the harm factor. In contrast, no indirect effects were found for the dependence factor. The direct effects of the consumption items as well as age and gender on the consequences are displayed in Table 5. Binge drinking increases substantially the probability to report any of the adverse consequences. With increasing frequency, feelings of guilt and concern of others are reported more often. With increasing quantity, blackouts become more frequent. However, the direct effects of frequency and quantity, although significant, are only marginal and much weaker than the effects of binging. Increased age is accompanied by a reduction of reports of role failure, guilt feelings, blackouts and injury, but not of loss of control and concern of others. The only direct effect of gender is apparent for guilt feelings, with men reporting less than women.

The associations between the three consumption items correspond to the continuous factorial structure obtained before: Frequency and quantity are only weakly associated with one another and with adverse consequences. However, both predict binging. Thus their strong indirect effect on the endorsement of consequences associated with the harm factor appears mediated through binging. With increased age both quantity and binging are reduced, but frequency of drinking is enhanced. Being male is associated with higher values on all three consumption items. Thus, gender apparently exerts its influence on adverse consequences primarily mediated by consumption. Taking into account significant covariate effects reduces not only the associations between factors and indicators, but also the correlation between the harm and the dependence factor from 0.84 (Table 4) to 0.61.

Table 6
Overall fit indicators for different class models.

	BIC[a]	nPar[b]	C_error[c]
LCA[d]			
1 class	26420	6	0%
2 classes	23110	13	5%
3 classes	23030	20	13%
4 classes	22996	27	13%
5 classes	23010	34	10%
LCFA[e]			
2 EFA[f] + covariates	22966	20	7%
Alcohol use[g]	20808	31	11%
Age and gender	20454	39	10%

[a] BIC = Bayes Information Criterium.
[b] nPar = number of model parameters.
[c] C_error = Classification error.
[d] LCA = Latent Class Analysis.
[e] LCFA = Latent Class Factor Analysis.
[f] EFA = Exploratory Factor Analysis.
[g] alcohol use: 3 consumption items.

3.4. *Latent class analyses to identify subgroups with distinct of AUDIT consequences (aim 4)*

Consumption items did not integrate satisfactorily into a common continuous factor structure. Therefore, we searched for subgroups based on the answer profiles only for the consequences items.

3.4.1. Number of latent classes. BIC values for successively tested LC models with an increasing number of nominal classes declined until a four class model with a CE of 13% (Table 6), to rise again for models with more than four classes. This suggests four subgroups with different adverse consequences profiles.

3.4.2. Dimensional ordering of latent classes. A LCFA model assuming four classes as ordered along two binary class dimensions obtained a lower BIC than the nominal four class model, and had a lower CE of 7% (Table 6). The BIC was further reduced when the consumption items together with age and gender were included as covariates in this LCFA model in the same way as before in the CFA models. However, their introduction did not improve, but slightly reduced classification precision.

Table 7 displays the logit coefficients reflecting the associations of the adverse consequences items with each of the two binary class factors. Multinomial regression relates a dichotomous dependent variable to dichotomous or polytomous independent variables. As in linear regression, the sign and size of coefficients in the model indicate the effect of a variable on the dependent variable. A positive coefficient represents the amount to which a unit change in the independent variable increases the log odds of being in the upper part as opposed to being in the lower part of the binary class factor. In order to provide more familiar, and easier to interpret

Table 7
Class–factor–indicator associations according to the latent class factor analysis model (linearly approximated standardized coefficients).

	Dependence (without covariate effects)	Harm (without covariate effects)
Loss control	0.56 (0.55)	0.26 (0.30)
Role failure	0.37 (0.42)	0.17 (0.19)
Guilt	0.38 (0.41)	0.51 (0.41)
Memory	0.29 (0.28)	0.39 (0.52)
Injury	0.21 (0.03)	0.48 (0.64)
Concern	0.31 (0.29)	0.41 (0.41)

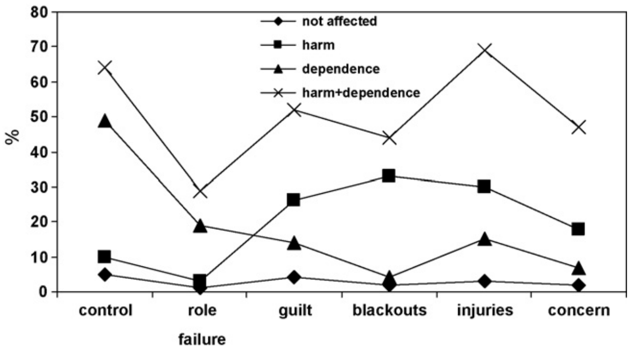

Fig. 2. Probability of endorsement of AUDIT adverse consequences in the four subgroups identified by LCFA.

FA output measures, Latent GOLD computes linearly approximated standardized maximum likelihood estimates of these coefficients (Table 7). In line with the dimensions identified in the continuous CFA, one binary class factor correlates strongest with loss of control and also substantially with role failure, while the second factor is primarily associated with the four harm items. Thus, the two dimensions emerging from the LCFA still have a focus on typical dependence symptoms on the one hand, and harmful consequences on the other hand, although these associations are less specific: role failure is associated more weakly with the first binary factor than in the CFA model, and the associations of this factor with feelings of guilt, blackouts, and concern of others are stronger. However, the two binary class factors are correlated markedly weaker (0.30), implying that the two discrete class factors are more clearly separated than the corresponding continuous dimensions.

3.4.3. Size and composition of AUDIT consequences subgroups. By cross-classifying the high and the low levels of the two class factors, four subgroups are defined (Fig. 2). According to the conditional class answer probabilities, in one class the probability to endorse any of the consequences items is maximally 5%. Without taking into account the influence of covariates (see below), 75% of the respondents were assigned to this unaffected subgroup. In a second subgroup comprising 10% of the respondents, the probability of harmful consequences varied between 18% and 33%, while the probability of loss of role function was 3% and of loss of control 10%. Thus, this subgroup is specifically characterized by endorsement of harmful consequences and will be labeled harm subgroup. In a third subgroup (13%), in contrast, loss of control and loss of role function are more than four times as likely to be affirmed than in the harm subgroup. However, the probability to endorse feelings of guilt and concern of others is only half of that for the harm subgroup. Thus, dependence symptoms are prominent in this subgroup, while harmful consequences are mostly missing (dependence subgroup). In a fourth subgroup with 12% of the respondents, the probabilities to endorse both harm and dependence consequences are high (harm and dependence subgroup). Guilt feelings is the most prominent adverse consequences, with a probability of endorsement surpassing 70%, but any of the remaining adverse consequences is also more likely to be endorsed in this subgroup than in any of the other subgroups.

Morning drinking, which was excluded from the LCFA because of its rare occurrence, fits well into the differential consequences profiles of the four subgroups: 6% of the members of the harm and dependence subgroup answer this item in the affirmative, but only 1% of the harm and the dependence subgroup, and no one of the unaffected subgroup.

3.5. Adjusting subgroups for differences in consumption, age and gender (aim 5)

Influences of these covariates on the class specific item profiles mirror those obtained for the continuous factors (Table 5). In addition the quantity of drinking now predicts concern of others. Thus we will not report these associations separately. Including the covariates into the LCFA model did generally reduce the indicator–factor associations (Table 7). This reduction results from adjusting for spurious covariation among the consequences items due to simultaneous effects of consumption, age and gender on consequences. The corresponding adjustment affected the subgroup composition (Table 8): The percentage of respondents with high values on the dependence dimension was increased from 12% to 19% and the percentage of respondents with high values on the harm factor from 15% to 25%. The unaffected subgroup shrank (from 75% to 66%) and the harm and dependence subgroup expanded (from 2% to 10%). The allocation rates to the other two subgroups remained quite stable (Table 8). Adjusting for differing, and in part opposite

Table 8
Class conditional endorsement probabilities [%] for AUDIT and DrInC adverse consequences in the four subgroups (class sizes [%]) and in the total sample.

	Unaffected 66[a] (75)[b]	Harm 15 (13)	Dependence 9 (10)	Harm + Dependence 10 (2)	Sample $N = 6259$
AUDIT consequence items					
Class indicators					
Loss of control	5	10	49	64	15
Role function	1	3	19	29	6
Blackouts	4	26	14	52	13
Injury	2	33	4	44	11
Guilt feelings	3	30	15	69	14
Concern	2	18	7	47	9
Morning drinking[c]	0	1	1	6	1
DrInC[d] items					
DUI[e]	4	10	6	14	1
Awkward	7	19	14	32	6
Sleep	8	18	14	28	12
Hangover	23	47	33	61	31
Vomited	5	10	10	17	7

[a] Class sizes with covariates (consumption items, age, gender).
[b] Class sizes without covariates in brackets.
[c] Not used as class indicator.
[d] Drinker Inventory of Consequences.
[e] Driving under the influence of alcohol.

influences of these covariates is essential to adequately judge the sizes and the composition of consequences subgroups.

3.6. *Differentiating the AUDIT subgroups according to further adverse drinking consequences (aim 6)*

The frequencies of endorsement of adverse consequences according to the DrInC in the total sample range from 1% (driving while under the influence of alcohol) to 31% (hangover). According to Table 8, any of the DrInC consequences occurred at least twice as often in any of the three problem subgroups than in the unaffected subgroup. Also, more members of the harm and dependence subgroup than of either the harm or the dependence subgroup endorsed each DrInC item. Finally, each DrInC item was affirmed more frequently in the harm than in the dependence subgroup, corroborating that these two subgroups differ with respect to the extent and severity of harmful consequences. Any of the four subgroups can be distinguished from any other by a distinctive pattern of adverse consequences which had not been used for the formation of the subgroups. Descriptively, this corroborates the external validity of the identified AUDIT consequences subgroups.

4. Discussion

Departing from a review of studies which generally suggested a two-dimensional structure of the answers to the AUDIT items, we had outlined six aims to be achieved by nonlinear factor and latent class analyses. The results obtained and their limitations will first be discussed in relation to these aims. Subsequently we will relate our dimensional results to the more general discussion of the dimensionality of AUD indicators.

Our first aim was to establish the dimensional structure of the AUDIT answers with statistical methods taking into account nonnormality of answer distributions and nonlinearity of the relations between indicators and constructs. The best fitting CFA model for the AUDIT answers of our sample of general practitioners' patients associated the consequences items either to a harm or to a dependence factor. The dependence factor emerged despite being based on only the two items loss of control and role failure, as the morning drinking item had to be dropped from the analyses because of its rare endorsement. The two dimensions were strongly associated, but statistically clearly separable, and well defined by differential loadings. In accordance with the intentions guiding the construction of the AUDIT (Saunders et al., 1993), the answers to the adverse consequences items indicate the influence of the two constructs harmful drinking and dependence. Thus, they do not indicate merely different degrees of severity of a general AUD dimension. Is this two-dimensional structure of the consequences indeed due to the use of appropriate nonlinear methods, or is it merely the consequence of special features of our sample? In the following, we consider several limitations which may impair the external validity of the two-dimensional structure of the consequences items.

First, we argued that the assumptions of normality and linearity of a dimensional structure underlying AUDIT answers are hardly ever met. Indeed, the distributions of the answers to the AUDIT items in our sample were exceedingly skewed (Table 2), as is to be expected in a sample not selected on the basis of alcohol problems. But even in samples composed of respondents fulfilling DSM criteria for dependence, items 6 (morning drinking) and 9 (injury to oneself or others) may still be negatively skewed, while items 1 and 10 will be positively skewed (Doyle et al., 2007). Thus, the assumption of normality will hardly ever hold for AUDIT data. Even when applying analysis methods appropriate for skewed data, it may be necessary to collapse answer categories to obtain sufficient observations for robust model estimations. In our analysis, all consequences items had to be dichotomized to yield sufficient answer category frequencies. However, as the use of dichotomized instead of full ordinal answer distributions does not impair the screening performance of the AUDIT (Bischof et al., 2007), dichotomizing does not alter the associations between indicators and constructs. Thus, our two-dimensional solution does not appear to result from a particular constellation of exceedingly skewed answer distributions.

Second, the prevalence of respondents meeting AUD and ARC criteria in our sample may have contributed to the two-dimensional solution. It is known that the variation of the prevalence of ARD and AUD among samples will influence the dimensional complexity of AUD indicators (Nelson et al., 1999). However, in the analysis of archival AUDIT data sets by Doyle et al. (2007), a three-dimensional solution did not fit worse than a two-dimensional solution in four data sets (Doyle et al., 2007; Table 3), including also two patient samples (Doyle et al., 2007; Table 2). Thus, even when based on continuous CFA models, three-dimensional solutions, also according to the study of Shevlin and Smith (2007), did rival two-dimensional ones across samples with widely varying prevalence of ARD and AUD. Only samples with very low prevalence of ARD and AUD may yield only one dimension for the AUDIT consequences (e.g., Gmel et al., 2001). While Gmel et al. (2001) report that about 9% of their respondents reach AUDIT sum scores ≥8, in our sample 15% are positive (9% women, 31% men). Although hazardous drinking is more frequent in our sample, the sample does not appear biased in this direction when compared to other surveys: A similar prevalence of AUDIT scores ≥8 (9% for women, 34% for men) was found in a representative survey of the German general population (Kraus and Augustin, 2005) and an even higher prevalence of 39% of the men and 19% of the women in a population survey in Scotland, England and Wales (Shevlin and Smith, 2007). We conclude from this comparison that the dimensional solution favored here might be obtained also in other samples than primary care patients, if the same methodology were applied.

Our second aim was to establish the functional status of the consumption items in a continuous factor model of the AUDIT item answers. Previous studies show that including all AUDIT items, and testing dimensions for the consumption and the consequences items, generally produces substantial associations of the consumption items with a proprietary factor, with the highest loading on the binging item (Doyle et al., 2007, Table 3). However, according to our analysis, treating the three consumption items as indicators of one latent variable does not account for plausible nonlinear relations among the three items: we found a negative association between the residuals for frequency and quantity, i.e., a second consumption factor besides a general consumption factor loading substantially on all three consumption items. On any level of consumption, drinking more frequently will tend to be associated with smaller quantities and vice versa. Thus, due to numerical and functional dependencies, the relations between the consumption items, and between them and the other AUDIT items are adequately captured neither by a two nor by a three-dimensional model. Previously, only Gmel et al. (2001) considered the difficulties associated with incorporating consumption information into one model together with latent variables for the consequences items. In their third and final model, they assigned each of the three consumption items to a unique latent variable. Apparently, the answers to the consumption items do not indicate the influence of a coherent, easily to interpret construct. Thus, as an alternative to modeling the consumption items as indicators of a construct, we treated them as descriptors of alcohol use behaviour by integrating them into the final two-dimensional model as covariates. The fit of this model was clearly better than that of models involving the consumption items as construct indicators. While quantity and frequency information influenced the position of respondents on the harm factor, binge drinking directly and substantially affected the answers to all consequences items. Thus, the information about frequency of binging is more salient for adverse consequences than frequency of drinking or typical amounts consumed. This is in line with other studies: Bergman and Källmen (2002) report that binging explained 34% of the variance of the AUDIT total score. In general, consumption information is associated more closely with the harm domain of consequences than with the dependence domain. As a consequence, screening for AUD by taking consumption information as proxy for all AUD, as done in brief AUDIT versions which rely exclusively on the first three AUDIT items (Reinert and Allen, 2007), will be less sensitive for signs of dependence than for signs of harmful drinking. Thus, the answer to the question whether adverse consequences are unidimensional or may be assigned to a dependence and a harm domain is important for attempts to improve screening for signs of the development of AUD.

Our third aim was to assess age and gender as predictors of consumption and consequences in this factorial model. Men consume more than women according to all three consumption items, but only one association of gender with consequences was found: Guilt feelings are rarer in men than in women. Increasing age is accompanied by an increase of the frequency of drinking, but a reduction in quantities consumed and in frequency of binging, and also a reduction of the probability of consequences. These results concur with findings from epidemiological studies (Day and Homish, 2002) and show that our sample is not different from those surveyed in other studies. However, the reported associations will influence results about the dimensional structure of the answers to the consequences items: Heterogeneity with respect to gender and age will increase the covariances among the consequences item answers, and thereby bias the dimensional analysis towards one factor. This is exemplified by the effect of introducing these covariates into our continuous factor model. In each case, introducing the covariates reduced the indicator–factor associations, but also the correlations between factors. This speaks again for including these variables as covariates in dimensional analyses of adverse consequences.

Our fourth aim was to probe for the existence of subgroups based on the consequences profiles. Latent class models were fitted to the adverse consequences to determine subgroups, and four subgroups were found. They were ordered along two discrete latent class factors which were differently associated with indicators of dependence and harm. This result corroborates the dimensional analyses: Two dimensions explain the answers to the consequences items better than one. The four subgroups were clearly differentiated also by their endorsement of consequences of drinking assessed by five of the DrInC items, corroborating their validity.

Our fifth aim was to delineate the influence of consumption as well as of age and gender on the allocation of respondents to subgroups. Including the consumption items as covariates together with age and gender led to a reduction of the size of the unaffected subgroup and an increase of the size of the harm and dependence subgroup.

Our sixth aim was to illustrate differences between the subgroups by reference to further adverse drinking consequences not used for the formation of the subgroups. The subgroups differed substantially in the probability of morning drinking (which had not been included in the LCFA because of its rarity) and five further consequences assessed with the DrInC. The groups appeared consistently ordered across all consequences, with probabilities of endorsement highest in the harm and dependence group, followed by the harm group, the dependence group and the unaffected group. Although only performed descriptively, this comparison supports the distinction between the subgroups.

In the following more general discussion, we try to reconcile our finding that two dimensions are necessary to explain the AUDIT adverse consequences with the results from studies which have investigated the dimensionality of other, not AUDIT-based, AUD indicators. The issue of the dimensional structure of the AUDIT corresponds to diverse results regarding the structure of DSM-IV alcohol abuse and dependence criteria in general population surveys. In some studies, evidence for two, although strongly associated factors was found (Muthén et al., 1993; Harford and Muthén, 2001). Hasin et al. (1994) and Proudfoot et al. (2006) also found two factors, but preferred a one factor model because of parsimony, as these two factors were correlated exceedingly high. Prior to an IRT analysis of DSM-IV criteria, Saha et al. (2006) established a unidimensional structure by an exploratory factor analysis of tetrachoric correlations. Preceding an IRT analysis of 110 alcohol-related problems assessed by interviewing middle aged men from the general population Krueger et al. (2004) also excluded a two-dimensional solution. According to these IRT analyses, DSM-IV criteria (Saha et al., 2006; Proudfoot et al., 2006) as well as alcohol-related problems (Krueger et al., 2004) appear ordered on one "severity" dimension, i.e., ordered with respect to the rarity of their occurrence. When searching for subgroups formed on the basis of DSM-IV criteria by means of LCFA Muthén (2006) found equivocal evidence for one and for two latent class variables, but retained only one factor to order the identified subgroups.

The necessity of two dimensions to explain adverse consequences as assessed by the AUDIT appears strikingly at odds with the proposed reformulation of DSM-IV criteria in preparation for the next version of the DSM. Differences between the just cited dimensional studies and our AUDIT analyses may be due to differences in the selection and formulation of items, but also to the different methods used to generate dimensional solutions. The factor analytic results of Saha et al. (2006) and Krueger et al. (2004) could not be biased by non-normality of the answer distributions, because they investigated tetrachoric correlations. However, they analyzed them with linear models which do not take into account nonlinear relations between indicators and constructs.

Despite our evidence for two adverse consequences dimensions, our results concur with the results of the mentioned IRT studies in several respects. According to Saha et al. (2006), loss of control (larger/longer) and loss of role function (neglect of roles) differed in severity but achieved the best discrimination. The authors concluded that the neglect of roles criterion should not have been moved from the DSM-III-R dependence category to the DSM-IV abuse category. The fact that both loss of role functioning and loss of control are explained by one latent variable in our study, both in the dimensional and in the class analysis, corroborates their stance. Second, only two of the three affected subgroups are characterized by a differentiation of consequences indicating either dependence or harmful drinking. In the harm and dependence subgroup all consequences are at least as likely to be affirmed as in the harm or the dependence subgroup. Third, the consequences assessed by the DrInC were clearly more likely in the overall affected subgroup than in the harm or the dependence subgroup.

Thus, in line with the IRT analyses, the LCFA of the AUDIT answers does not support a categorical diagnostic model which assigns AUD to either a dependence or an abuse category. On the other hand, according to the LCFA model the groups cannot be ordered sufficiently along only one severity dimension without substantial loss of information. Generalizing from the few adverse consequences assessed by the AUDIT, alcohol-related problems appear to vary along two dimensions in the medium range of severity, but they concur in the harm and dependence subgroup.

To conclude: In most situations in which screening for AUD is performed, one has to deal with heavily skewed distributions of answers to items which ask for respective signs, symptoms, or risk factors. When applied to these data, conventional psychometric analyses run the risk of producing heavily biased results. Recently, Midanik et al. (2007) have therefore called for the application of methods which do not rely on the assumptions of normality and Linearity in the alcohol field. We have shown for the AUDIT that such methods indeed unravel different dimensions than were found by linear methods. It may add to its credibility that the now proposed structure and the corresponding subgroups agree well with the original dimensional conception of AUD consequences which drove the development of the AUDIT.

Role of funding source

This research was supported by a grant from the Federal Ministry of Education and Research (No. 01EB0131). The ministry was not involved in the design of the study; the collection, analysis, and interpretation of the data; the writing of the report; or the decision to submit the paper for publication.

Contributors

Fred Rist and Ralf Demmel were responsible for the study plan. Angelika Glöckner-Rist performed the systematic review of the AUDIT dimensional analyses studies and conceived and conducted the structure and class analyses. Fred Rist and Angelika Glöckner-Rist drafted the manuscript. Ralf Demmel coordinated the study and helped to draft the manuscript. All authors read and approved the final manuscript.

Conflict of interest

The authors declare that they have no competing interests.

Acknowledgements

The authors thank Ines Aulhorn, Jutta Hagen and Barbara Scheuren for their diligent and reliable conduct of the data collection and data entry during the study period, Verena Zimmer for the careful preparation of the final data set, and Alwina Stein for help with the editing of the manuscript. Stefan Gesenhues has helped us considerably to recruit general practitioners for participation in the study.

References

Allen, J.P., Litten, R.Z., Fertig, J.B., Babor, T., 1997. A review of research on the Alcohol Use Disorders Identification Test (AUDIT). Alcohol Clin Exp. Res. 21, 613–619.

Babor, T.F., Higgins-Biddle, J.C., Saunders, J.B., Monteiro, M.G., 2001. AUDIT The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care, second ed. World Health Organization, Geneva.

Bergman, H., Källmen, H., 2002. Alcohol use among Swedes and a psychometric evaluation of the Alcohol Use Disorders Identification Test. Alcohol Alcohol. 37, 245–251.

Bischof, G., Grothues, J., Reinhardt, S., John, U., Meyer, C., Ulbricht, S., Rumpf, H.-J., 2007. Alcohol screening in general practices using the AUDIT: How many response categories are necessary? Eur. Addict. Res. 13, 25–30.

Bock, G.D., Gibbons, R., Muraki, E., 1988. Full information item factor analysis. Appl. Psychol. Meas. 12, 261–280.

Bollen, K.A., 1989. Structural Equations with Latent Variables. John Wiley & Sons, Oxford.

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. J. Clin. Psychiatry 64, 767–774.

Chin, W.W., 1998. Issues and opinion on structural equation modelling [Commentary]. MIS Q. 22, vii–xvi.

Chung, T., Colby, S.M., Barnett, N.P., Monti, P.M., 2002. Alcohol Use Disorders Identification Test: Factor structure in an adolescent emergency department sample. Alcohol. Clin. Exp. Res. 26, 223–231.

Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. Psychometrika 16, 297–334.

Cronbach, L.J., Shavelson, R.J., 2004. My current thoughts on coefficient alpha and successor procedures. Educ. Psychol. Meas. 64, 391–418.

Day, N.L., Homish, G.G., 2002. The epidemiology of alcohol use, abuse, and dependence. In: Tsuang, M.T., Tohen, M. (Eds.), Textbook in Psychiatric Epidemiology. John Wiley & Sons, Inc., New York, NY, pp. 459–477.

Demmel, R., Rist, F., Hagen, J., Aulhorn, I., Scheuren, B., Scherbaum, N., Gesenhues, S., Rollnick, S., 2003. Sekundärprävention – mehr als Screening und gute Ratschläge. Suchtmedizin in Forschung und Praxis 5, 33–36.

DeShon, R.P., 1998. A cautionary note on measurement error corrections in structural equation models. Psychol. Methods 3, 412–423.

Doyle, S.R., Donovan, D.M., Kivlahan, D.R., 2007. The factor structure of the Alcohol Use Disorders Identification Test (AUDIT). J. Stud. Alcohol Drugs. 68, 474–479.

El-Bassel, N., Schilling, R., Ivanoff, A., Chen, D.-R., Hanson, M., 1998. Assessing the World Health Organization's Alcohol Use Disorder Identification Test among incarcerated women. J. Offender Rehab. 26, 71–89.

Glöckner-Rist, A., Hoijtink, H., 2003. The best of both worlds: Factor analysis of dichotomous data using item response theory and structural equation modeling. Struct. Equation Model. 10, 544–565.

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 Alcohol Depend. 64, 151–163.

Gordon, A.J., 2006. Screening the drinking: Identifying problem alcohol consumption in primary care settings. Adv. Stud. Med. 6, 137–147.

Hagenaars, J.A., McCutcheon, A.L., 2006. Applied Latent Class Analysis. Cambridge University Press, Cambridge.

Harford, 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. J. Stud. Alcohol 62, 150–157.

Hasin, D.S., Muthuen, B., Wisnicki, K.S., Grant, B., 1994. Validity of the bi-axial dependence concept: A test in the US general population. Addiction 89, 573–579.

Heinen, T., 1996. Latent class and discrete latent trait models: Similarities and differences. Sage Publications, Inc., Thousand Oaks, CA.

Karno, M., Granholm, E., Lin, A., 2000. Factor structure of the Alcohol Use Disorders Identification Test (AUDIT) in a mental health clinic sample. J. Stud. Alcohol 61, 751–758.

Kelly, T.M., Donovan, J.E., 2001. Confirmatory factor analyses of the Alcohol Use Disorders Identification Test (AUDIT) among adolescents treated in emergency departments. J. Stud. Alcohol 62, 838–842.

Kraus, L., Augustin, R., 2005. Alkoholkonsum, alkoholbezogene Probleme und Trends. Ergebnisse des Epidemiologischen Suchtsurvey 2003. Sucht. 51, 29–39 (special issue).

Krueger, R.F., Nichol, P.E., Hicks, B.M., Markon, K.E., Patrick, C.J., Iacono, W.G., et al., 2004. Using latent trait modeling to conceptualize an alcohol problems continuum. Psychol Assess. 16, 107–119.

Laumeyer, S., 2003. Psychometrische Eigenschaften und Faktorenstruktur einer deutschsprachigen Version des Drinker Inventory of Consequences [A German-language version of the Drinker Inventory of Consequences]. Unpublished master's thesis, University of Münster, Münster, Germany.

Lima, C.T., Freire, A.C.C., Silva, A.P.B., Teixeira, R.M., Farrell, M., Prince, M., 2005. Concurrent and construct validity of the AUDIT in an urban Brazilian sample. Alcohol Alcohol. 40, 584–589.

Magidson, J., Vermunt, J.K., 2001. Latent class factor and cluster models, bi-plots, and related graphical displays. Sociol. Methodol. 31, 223–264.

Maisto, S.A., Conigliaro, J., McNeil, M., Kraemer, K., Kelley, M.E., 2000. An empirical investigation of the factor structure of the AUDIT. Psychol. Assess. 12, 346–353.

McDonald, R.P., 1996. Normal ogive multidimensional model. In: van der Linden, W.J., Hambleton, R.K. (Eds.), Handbook of Modern Item Response Theory. Springer, New York, pp. 258–270.

Medina-Mora, E., Carreno, S., de la Fuente, J.R., 1998. Experience with the Alcohol Use Disorders Identification Test (AUDIT) in Mexico. Recent Dev. Alcohol. 14, 383–396.

Midanik, L.T., Greenfield, T.K., Bond, J., 2007. Addiction sciences and its psychometrics: The measurement of alcohol-related problems. Addiction 102, 1701–1710.

Miller, W.R., Tonigan, J.S., Longabaugh, R., 1995. The Drinker Inventory of Consequences (DrInC). An instrument for Assessing Adverse Consequences of Alcohol Abuse. National Institute on Alcohol Abuse and Alcoholism Project MATCH Monograph Series, vol. 4. National Institutes of Health, Rockville, MD.

Mislevy, R.J., 1986. Bayes model estimation in item response models. Psychometrika 51, 177–195.

Muthén, B.O., 2002. Beyond SEM: General latent variable modeling. Behaviormetrika 29, 81–117.

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

Muthén, B.O., Grant, B., Hasin, D., 1993. The dimensionality of alcohol abuse and dependence: Factor analysis of DSM-III-R and proposed DSM-IV criteria in the 1988 National Health Interview Survey. Addiction 88, 1079–1090.

Nelson, C.B., Rehm, J., Üstün, 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.

O'Hare, T., Sherrer, M.V., 1999. Validating the Alcohol Use Disorder Identification Test with college first-offenders. J. Subst. Abuse Treat. 17, 113–119.

Proudfoot, H., Baillie, A.J., Teesson, M., 2006. The structure of alcohol dependence in the community. Drug Alcohol Depend. 81, 21–26.

Reckase, M.D., 1996. A linear logistic multidimensional model for dichotomous item response data. In: van der Linden, W.J., Hambleton, R.K. (Eds.), Handbook of Modern Item Response Theory. Springer, New York, pp. 271–286.

Reinert, D.F., Allen, J.P., 2007. The Alcohol Use Disorders Identification Test: An Update of Research Findings. Alcohol. Clin. Exp. Res. 31, 185–199.

Rist, F., Scheuren, B., Demmel, R., Hagen, J., Aulhorn, I., 2003. Der Münsteraner Alcohol Use Disorders Identification Test (AUDIT-G-M). In: Glöckner-Rist, A., Rist, F., Küfner, H. (Eds.), Elektronisches Handbuch zu Erhebungsinstrumenten im Suchtbereich, Version 3.00. Zentrum für Umfragen, Methoden und Analysen e. V, Mannheim.

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. Psychol. Med. 36, 931–941.

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.

Shevlin, M., Smith, G.W., 2007. The factor structure and concurrent validity of the Alcohol Use Disorder Identification Test based on a nationally representative UK sample. Alcohol Alcohol. 42, 582–587.

Shields, A.L., Guttmannova, K., Caruso, J.C., 2004. An Examination of the factor structure of the Alcohol Use Disorders Identification Test in two high-risk samples. Subst. Use Misuse 39, 1161–1182.

Skipsey, K., Burleson, J.A., Kranzler, H.R., 1997. Utility of the AUDIT for identification of hazardous or harmful drinking in drug dependent patients. Drug Alcohol Depend. 45, 157–163.

Takane, Y., DeLeeuw, J., 1987. On the relationship between item response theory and factor analysis of discretized variables. Psychometrika 52, 393–408.

Tomarken, A.J., Waller, N.G., 2005. Structural equation modelling: strengths, limitations, and misconceptions. Annu. Rev. Clin. Psychol. 1, 31–65.

Vermunt, J.K., Magidson, J., 2004. Factor analysis with categorical indicators: a comparison between traditional and latent class approaches. In: Van der Ark, L.A., Croon, M.A., Sijtsma, K. (Eds.), New Developments in Categorical Data Analysis for the Social and Behavioral Sciences. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 41–62.