Alcohol & Alcoholism Vol. 41, Supplement 1, pp. i19–i25, 2006
doi:10.1093/alcalc/agl072, available online at www.alcalc.oxfordjournals.org

A COMPARISON OF THE ALCOHOL USE DISORDER IDENTIFICATION TEST (AUDIT) IN GENERAL POPULATION SURVEYS IN NINE EUROPEAN COUNTRIES

RONALD A. KNIBBE[1]*, MIEKE DERICKX[1], SANDRA KUNTSCHE[2], ULRIKE GRITTNER[3] and KIM BLOOMFIELD[3,4]

[1]University Maastricht, Health Care Science, Section Medical Sociology, PO Box 616, 6200 MD Maastricht, [2]Swiss Institute for the Prevention of Alcohol and Drug Problems, Lausanne, Switzerland, [3]Institute for Biostatistics and Clinical Epidemiology, Charité—University Medicine Berlin, Germany and [4]Unit for Health Promotion Research, University of Southern Denmark, Esbjerg, Denmark

Abstract — Aims: This study explored the suitability of the Alcohol Use Disorder Identification Test (AUDIT) for cross-national comparable estimates of problem drinking in general populations. On the item level the focus is on responsiveness to cross-national and gender differences. For the set of items the focus is on intercorrelations between items, indicating to what extent the AUDIT constitutes a scale. **Methods:** General population surveys from nine European countries were included. Cross-tabulations were used to analyse cross-national and gender differences in scores on the items. Reliability analysis was used to analyse intercorrelations between the items. **Results:** The items 'blackouts' (men and women) and 'guilt and remorse' (women) are the most frequently reported consequences. Gender differences tended to be smaller for 'guilt and remorse' and 'concern of others', and largest for 'morning drinking'. The reliability analysis shows that in eight of the nine countries frequency of drinking lowers the alpha. Injury and concern of others lead to a lower internal consistency in three countries. **Conclusions:** There was sufficient variation between countries in the pattern of responses and variation in gender differences to conclude that the set of consequence items was responsive to national and gender differences in problem drinking. Frequency of drinking was not a good indicator of problem drinking. The country differences in item total correlations of consequences might be due to differences in how these items are interpreted. Decisions on which items to include in an instrument to allow comparison of estimates of problem drinking cross-nationally require studies on how these items are interpreted in general populations of different countries.

INTRODUCTION

Prevalence estimates of problem drinking are mostly made on the basis of self-reports of respondents in general population surveys. Biomedical markers for alcohol misuse are restricted to the time period during which alcohol misuse can be detected, and cover only a limited range of alcohol-related problems (Beresford *et al*., 1990; Conigrave *et al*., 1995). Generally speaking, two strategies have been used to measure problem drinking in population surveys. The first focuses on the variety of problems possibly due to alcohol consumption and aims to establish how often these occur (e.g. Cahalan, 1976). The advantage of a more detailed measurement of alcohol-related problems is that it provides a more comprehensive overview of the types of problems associated with alcohol consumption. This allows, for example, specifying which types of alcohol-related problems dominate in a particular subpopulation, region, or country; the disadvantage is, however, that this requires many questions. The second strategy is to reduce the range of consequences to a limited set of items which allows establishing whether or not a person is a problem drinker. The advantage of a limited set of items to screen for problem drinking is its brevity, whereas the disadvantage is that such an instrument is limited in its ability to take into account the variability in the types of problems associated with alcohol consumption.

This variability is important both within and between countries. For example, within some subpopulations or countries drunkenness, violence, and accidents may be the most prevalent types of alcohol-related problems; in other subpopulations, chronic health consequences due to excessive consumption and work-related problems may be the most prevalent problems. Short screening instruments pre-suppose uniformity in alcohol-related problems within general populations. Regarding differences in drinking patterns, there may be more diversity in types of alcohol-related problems within countries (e.g. between men and women, younger, and older persons, social classes, etc.) and especially between countries (e.g. Mediterranean countries with a higher daily consumption of wine, and Scandinavian countries with a higher frequency of risky single-occasion drinking) than a short screening instrument is able to cover.

Within the European context, several countries have a (more or less) established tradition concerning research on alcohol consumption and alcohol-related problems. However, these traditions differ in items measuring alcohol consumption and alcohol-related problems as well as in how to combine items into variables indicating more specific concepts (e.g. dependence, misuse, and problem drinking) or also a general concept such as 'problem drinker' (Knibbe *et al*., 2003). This variety may not be a problem if evaluated from the limited perspective of each particular study. Thus, the comparability with outcomes from other studies in the same country may be seen as more important than the comparability with studies from other countries. However, there is an increasing demand for comparable prevalence estimates from different countries within Europe, possibly due to the increasing number of countries within the EU. Perhaps more important is the increased recognition on the European level that alcoholic beverages have not only an economic aspect but also an impact on public health (European Commission, 2006). To ascertain the public health aspects of alcohol consumption at the European level, comparable estimates of alcohol misuse and problem drinking are required.

*Author to whom correspondence should be addressed at: Tel: +31 43 3882204; Fax: +31 43 3884169; E-mail: r.knibbe@zw.unimaas.nl

Therefore, this study has explored whether the Alcohol Use Disorder Identification Test (AUDIT) (Saunders *et al*., 1993a) could be an appropriate instrument to provide comparable estimates of problem drinking in general populations from different European countries. General population surveys from nine European countries are used in which all or most of the 10 items of the AUDIT were measured in a comparable way.

Compared with other well-known screening instruments like CAGE (Mayfield *et al*., 1974), SAAST (Davis *et al*., 1987), and SMAST (Selzer *et al*., 1975), the most distinguishing characteristics of the AUDIT are as follows:

(i) It has been developed to detect problem drinking in a general treatment setting; several countries were involved in the development of the AUDIT.
(ii) Conceptually, the AUDIT is assumed to have three different domains: quantity and frequency of drinking (with three items: frequency, and quantity of drinking, and frequency of six or more glasses); dependence (three items: unable to stop, failing normative expectations, and morning drinking), and alcohol-related harm (four items: guilt, blackout, injury, and concern of others). When following the DSM-IV diagnostics blackout is used as an indicator of dependence while failing normative expectations is an indicator of alcohol abuse rather than an indicator of dependence.

The usefulness of the AUDIT in different national and cultural contexts was an important issue in the development of the instrument. From that viewpoint, the AUDIT has been well applied in various countries (Babor *et al*., 2001). However, the AUDIT was not developed to provide prevalence estimates in general populations and, with a few exceptions (Saunders *et al*., 1993b; Holmila, 1995; Fleming, 1996; Medina-Mora *et al*., 1998; Ivis *et al*., 2000), has not been used for that purpose. When using the AUDIT to estimate prevalence rather than for (early) detection in a treatment setting the first thing to mention is, of course, that identifying an anonymous respondent in a survey with the AUDIT is practically and conceptually quite different from using the instrument in a treatment context with the intention of providing a brief intervention. Practically, it is different because prevalence estimates are used in the context of policy rather than treatment; conceptually, because qualifying for intervention assumes a diagnosis for which survey methods alone are rarely, if ever, sufficient. There are also three more specific points which need to be considered when using the AUDIT in the context of surveys to estimate the prevalence. First, the interpretation of responses to the AUDIT items is more controlled in a treatment setting than in general population surveys. Kypri *et al*. (2002) demonstrated that for two items (i.e. failure to fulfil role expectations, and alcohol-related injury) students are likely to interpret these differently than originally intended. Second, compared with populations entering a treatment setting, general populations are likely to be much more heterogeneous in their drinking pattern and the related consequences due to alcohol consumption. This is important considering that the scores of individual items are summed to establish a sum score. Third, the conceptually different dimensions constituting the AUDIT may mean that a psychometric analysis would show low intercorrelations between all or part of the items of the whole set. However, low intercorrelations between all or some of the items would mean that a sum score based on the whole set of items is less meaningful. Just as important in this context is that countries may differ in intercorrelations between the items. In that case sum scores mean different things in different countries and are not really comparable. Both the greater heterogeneity in general population samples and the distinct dimensions of the AUDIT may influence the extent to which the set of items constitutes an internally consistent scale.

In a Swiss general population sample, Gmel *et al*. (2001) found that the indicators of drinking pattern, and especially frequency of drinking, show almost no correlation with the items measuring dependence or harmful consequences. They conclude that frequency of drinking should not be included in a screening instrument for alcohol-related problems in the Swiss general population (of course this may not apply to other countries). For the alcohol-related problems, Gmel *et al*. (2001) could not find a two-factor (dependence and alcohol-related harm) structure; rather, all these items seemed to indicate the same underlying factor. However, there may be relevant differences between countries regarding the extent to which all these items form an internally consistent scale and, consequently, there may also be differences between countries in how the sum scores should be interpreted.

In this study we first inspect differences between countries on the item level. Such differences give an initial impression of national and cultural variability in the specific type of consequence most likely to be associated with alcohol misuse. Therefore, such variations give a preliminary indication of the sensitivity of the AUDIT for (sub)cultural differences in problems associated with drinking. A special focus of this paper is placed on gender differences within countries at the item level; it is known that men and women may differ in the specific problems that alcohol use may lead to. In most studies, because women drink less than men, such differences would not appear as a higher prevalence in some consequences among women. A better indication of the gender sensitivity of the set of items is the gender ratio. Assuming that men and women differ in how they drink and the consequences that are likely to result from drinking, the variation in gender ratio would give an indication of the extent to which the items also cover problem drinking among women.

To evaluate the entire set of items we analysed the extent to which the items of the AUDIT form an internally consistent scale in each country; differences found between countries imply that the sum scores of the AUDIT have to be interpreted differently. From a statistical viewpoint, in countries in which the items together constitute an internally consistent scale one may assume that responses can be combined to rank order drinkers according to severity of problems. However, for those countries in which the summed items do not form an internally consistent scale the score on the set of items does not indicate (or less reliably indicates) differences in the severity of alcohol problems. Of course, differences between countries in the extent to which the items taken together are a statistically reliable scale also influence the comparability of AUDIT scores between countries. Special attention has been given to the contribution of the drinking indicators to the internal consistency of the set of items.

The following research problems guided the analysis of the items in the AUDIT:

(i) Do differences exist between countries on the items constituting the AUDIT, and do gender differences exist within countries on these items?
(ii) Do countries differ in the extent to which the set of items constitute an internally consistent scale?
(iii) Do countries differ in how the drinking indicators used in the AUDIT contribute to the internal consistency of the AUDIT?

DATA AND METHODS

This study is based on data from the broader GENACIS project. The analyses were limited to nine European countries from the present EU concerted action and the broader GENACIS project. With the exception of Spain and The Netherlands all other surveys were national surveys. All surveys included all or most of the AUDIT-items in their questionnaire. Further details concerning the data can be found in the introduction to this issue. The article by Mäkelä *et al* in this issue also provides information on construction of drinking variables.

The age range is similar between countries: the youngest age category is from 15 (e.g. Switzerland) to 19 years (e.g. Hungary), and the oldest 64 years (Czech Republic) or older (all other countries). All surveys include both sexes, and the surveys took place between 1997 and 2002 in the nine countries. Table 1 shows that because countries differ in their mode of interview and non-response rates, the surveys may differ in underreporting of the consequences (e.g. more underreporting in face-to-face interviews than in postal interviews) and in the extent of selective non-response among heavier drinkers. These issues will not be discussed here, except to emphasize that direct comparisons of prevalences between countries should be made with caution. For the purpose of this study (i.e. a cross-national comparison of scores on the items and the intercorrelation between items), the main point is whether these surveys cover most of the variations in drinking pattern and the consequences of these populations. We assume that in this respect all surveys included here are adequate.

The AUDIT consists of 10 questions (Table 2). A first distinction can be made between the first three items measuring alcohol use and the seven other items measuring alcohol-related consequences.

Within the set of alcohol-related consequences a further distinction can be made between items 4, 5, and 6 measuring dependence and items 7–10 measuring alcohol-related harm. Of the nine countries included in this study, six countries included all seven items on consequences, Iceland included six items, and the UK and The Netherlands each included five of the consequence items (see Table 3). The consequences of drinking were established by asking direct questions (e.g. how often during the past year were you not able to stop drinking once you had started?). In all countries (except The Netherlands) these items had five answer categories, ranging from 'never' to 'daily or almost daily'; however, there were considerable differences in the precise cut-off points used in each country. To make the answers to these items comparable we dichotomized the answers into 'never' or 'at least once in the past 12 months'. There were also slight differences in the wording of the questions and the answer categories used to measure the drinking indicators. However, despite these differences it was possible to construct variables which were comparable over countries. When the AUDIT is used for detection in a treatment setting all items about alcohol-related consequences (except 'concern of others' and 'injury') are scored on a scale from 0 to 4. 'Concern of others' and 'injury' score 0 (never), 2 (yes, but not in the past year), and 3 (yes, during the past year). A sum score of 8 (men) or 7 (women) is indicative of hazardous or harmful drinking; a score of 13 or higher is indicative for alcohol-related harm. However, as mentioned above, to increase comparability, we simplified all answers to the questions addressing consequences by creating only two response categories (i.e. never/at least once in the past 12 months). For the drinking indicators we followed the guidelines of the AUDIT in using a cut-off point of drinking twice a week or more often for frequency, consuming five or more drinks for quantity per occasion, and consuming six or more drinks once a month or more often.

In all analyses, abstainers (defined as not having consumed alcoholic beverages in the past 12 months) were excluded. The data were analysed with SPSS 11.0 (SPSS, 2002). Crosstabulations were used when comparing prevalence on the level of items. To analyse intercorrelations between items, reliability analysis was used. Cronbach's alpha was used as an overall measure to analyse intercorrelations. A standardized alpha requires a similar variance on the item level; because this condition was not met (see Table 3), the alpha (rather than the standardized item alpha) was used as an indicator for intercorrelations.

Table 1. Survey characteristics

Country	Year	Sampling frame	Age range (years)	Survey mode	Response rate
Switzerland	1997	National	15+	Telephone	68.4%
Spain	2003	Regional	18+	Face to face (sensitive questions self-administered)	Quota
UK	2000	National	18+	Face to face and CAPI	Quota
Sweden	2002	National	17–82	Telephone	67.8%
Finland	2000	National	16–70	Face to face (AUDIT self-administered)	79.4%
The Netherlands	1999	Regional	16–69	Postal	71.0%
Czech Republic	2002	National	18–64	Face to face	72.6%
Hungary	2001	National	19–65	Face to face (alcohol questions self-administered)	Quota
Iceland	2001	National	18–75	Mixed (half/half postal and telephone survey)	70.1% or 56.6%

Table 2. Items of the Alcohol Use Disorder Identification Test (AUDIT)

No.	Item
(1) Frequency	How often do you have a drink containing alcohol?
(2) GL. per drinking day	How many drinks containing alcohol do you have on a typical day when you are drinking?
(3) Frequency 6+ Drinking	How often do you have six or more drinks on one occasion?
(4) Unable to stop	How often during the past year have you found that you were not able to stop drinking once you had started?
(5) Failed normative expectations	How often during the past year have you failed to do what was normally expected of you because of drinking?
(6) Morning drinking	How often during the past year have you needed a first drink in the morning to get yourself going after a heavy drinking session?
(7) Guilt or remorse	How often during the past year have you had a feeling of guilt or remorse after drinking?
(8) Blackout	How often during the past year have you been unable to remember what happened the night before because of your drinking?
(9) Injury	Have you or someone else been injured because of your drinking?
(10) Concern of others	Has a relative, friend, doctor, or other health care worker been concerned about your drinking or suggested you cut down?

RESULTS

Consequences measured with the AUDIT

Table 3 presents the proportion of men and women and the gender ratio in each country that reported on consequences due to alcohol use. For men, in six of the nine countries blackout is the most frequently reported consequence. In Switzerland, Hungary, and Iceland it is the second most frequently reported consequence after normative expectations (Switzerland), morning drinking (Hungary), and guilt/remorse (Iceland). Although there is some uniformity between countries as to which consequence men are most likely to report, differences in the prevalence of each of the consequences are large. For example, for blackouts the prevalence among men ranges from 42.6% (Finland) and 30.8% (Czech Republic) to 8.7% (Hungary) and 7.6% (Switzerland). For women there is more variability in the items most frequently reported as a consequence. Blackouts are most frequently reported by women in five countries (UK, Sweden, The Netherlands, Spain, and Iceland), whereas in Finland and the Czech Republic it is the second most frequently reported. Guilt or remorse is the most frequently reported consequence by women in Finland and the Czech Republic and the second in Spain, UK, Sweden and Hungary.

In six of the nine countries the gender difference tends to be comparatively smaller (<1.5) for guilt and remorse. In three of the seven countries which included 'concern of others' the gender ratio approaches 1, i.e. Spain, Czech Republic, and Hungary. For 'morning drinking' the gender ratio tends to be highest, ranging from 1.5 (Spain) to 8.0 (Sweden). In most countries (almost) all items show gender differences of 2 or higher. The exceptions are Spain, Iceland, and The Netherlands where only one of the items had a gender ratio higher than 2. Gender differences tend to be highest in Switzerland (2.3-7.0) and in Hungary (2.4-4.3).

AUDIT indicators for drinking

Table 4 presents the prevalence estimates for positive responses to each AUDIT item.

In all countries except Sweden, the highest percentage is found for frequency of drinking. In Switzerland the percentage of men and women drinking $\geqslant$2 times a week is the highest. The lowest percentages of drinking $\geqslant$2 times a week are in Sweden and (for women only) in Hungary.

Consuming five or more drinks per occasion (quantity/occasion) is most often reported by men and women in the Czech Republic (60.2% and 24.9%, respectively) followed by Iceland (men: 33.1%; women: 21.3%) and Finland (men: 39.4%; women: 16.9%). For consuming six or more drinks per day the highest percentages are found for men and women in Finland, Iceland and Sweden. In all countries gender differences are lowest for frequency of drinking. In Hungary comparatively high gender ratios are found for all drinking indicators. Compared with the consequence items (Table 3), much higher percentages scored on the drinking indicators (Table 4). This indicates that the consequence items select more specific categories of drinkers than do the drinking indicators.

Correlations between items of the AUDIT

Table 5 shows the item total correlations and two Cronbach's alpha coefficients for each country: one computed with consequences only, and one computed with both consequences and drinking indicators. Table 5 also presents the worst items from a statistical viewpoint; the criterion was that exclusion of these items would not change or even decrease Cronbach's alpha.

Of all the items, frequency of drinking has the lowest item total correlation for drinking items in all countries except Hungary. Of the consequence items, injury has the lowest item total correlation in five of the eight countries which have this item. In a sixth country, Hungary, the item total correlation of injury (0.12) is also very low; however, in Hungary, the item total correlation of concern of others is even lower and negative (-0.05). In the Czech Republic the item total correlation of concern of others is also very low (0.05).

Inclusion of the drinking indicators in the analysis leads to a lower Cronbach's alpha in three countries (Switzerland, Spain, and UK). In one country (Hungary) Cronbach's alpha does not change and in five countries (Sweden, Finland, The Netherlands, Czech Republic, and Iceland) the alpha only marginally increases. Inspection of how each item contributes to Cronbach's alpha shows that in all countries (except the Czech Republic), the alpha would improve when frequency of drinking is left out. Of the consequence items, 'injury due to drinking' decreases the alpha in Switzerland, Spain and Hungary. 'Concern of others' decreases the alpha in the Czech Republic and Hungary. In the UK and Iceland the alpha is decreased by 'morning drinking' and 'guilt/remorse', respectively.

Cronbach's alpha, computed with the consequence items only, is lower in Switzerland (0.61) and The Netherlands (0.59) compared to all other countries (0.68-0.78). For The Netherlands this is mostly due to the smaller number of items.

Table 3. Prevalence estimates (%) and gender ratios for AUDIT indicators of consequences

	Switzerland			Spain			UK			Sweden			Finland			The Netherlands			Czech Republic			Hungary			Iceland		
	M	F	R	M	F	R	M	F	R	M	F	R	M	F	R	M	F	R	M	F	R	M	F	R	M	F	R
Unable to stop	5.9	2.2	2.7	5.2	4.7	1.1	12.4	6.4	1.9	6.8	4.0	1.7	16.0	8.5	1.9	6.4	6.6	1.0	8.9	4.4	2.0	6.0	2.4	2.5	14.1	7.3	1.9
Normative expect	8.0	2.7	3.0	7.0	5.3	1.3	16.5	8.1	2.0	12.0	5.2	2.3	13.7	6.9	2.0				26.8	11.7	2.3	6.0	1.8	3.3	15.9	8.9	1.8
Morning drinking	1.4	0.2	7.0	2.6	1.7	1.5	4.0	1.0	4.0	2.6	0.4	6.5	14.8	3.2	4.6				17.6	4.7	3.7	8.9	2.2	4.1	11.0	2.4	4.6
Guilt/remorse	6.4	2.3	2.8	7.5	6.5	1.2	16.5	11.1	1.5	9.9	6.8	1.5	39.4	25.9	1.5	9.3	7.1	1.3	29.9	18.9	1.6	6.6	2.3	2.9	26.4	19.3	1.4
Blackout	7.6	1.8	4.2	10.4	7.7	1.4	25.7	12.8	2.0	18.0	9.2	2.0	42.6	18.0	2.4	16.5	9.3	1.8	30.8	14.3	2.2	8.7	3.0	2.9	22.1	13.4	1.7
Injured	1.6	0.7	2.3	4.2	1.3	3.2				3.3	1.4	2.4	21.5	10.7	2.0	1.9	2.3	0.8	19.7	6.1	3.2	0.9	0.4	2.3	9.2	4.9	1.9
Concern of others	5.6	2.1	2.7	5.1	4.3	1.2				4.7	2.1	2.2	17.6	5.6	3.1	8.8	3.2	2.8	10.6	11.8	0.90	5.9	6.1	0.97			

M = male; F = female; R = gender ratio.

When Cronbach's alpha is computed for the same selection of consequences for countries other than The Netherlands, the differences between countries in the alpha are less than 0.05 (results not presented). For Switzerland it can be concluded that, compared with other European countries, the interrelations between the consequence items are lower than in all other countries.

DISCUSSION

The main purpose of this study on the AUDIT is to inspect on the item level the differences between countries and gender differences and the intercorrelations between all items. National and gender differences on the item level are interpreted in the context of responsiveness of the AUDIT to national and gender differences. The intercorrelations are interpreted in the context of how well the sum score of the AUDIT rank orders drinkers in each of the countries according to severity of drinking problems.

On the level of the individual consequence items, the countries do not differ greatly regarding which consequence is most likely to be reported. For men, in most countries this is 'having blackouts'. Among women 'blackouts' and 'guilt and remorse' are most frequently reported. Although differences as to which consequence is most likely to be reported are not large, the countries differ considerably in the percentages reporting consequences. In Finland and the Czech Republic the percentages of men and women reporting consequences tend to be highest while in Switzerland, Spain, and Hungary smaller percentages of men and women report consequences. Within countries there is considerable variation in gender ratios. For the drinking indicators percentages much higher than the cut-off points were found, especially for frequency of drinking twice a week or more often. However, also for these indicators there are substantial differences between countries and also for gender differences within countries. On the item level there appears to be sufficient variation between countries in the pattern of responses to the AUDIT items and in gender differences, to conclude that this set of items is responsive to national and gender differences in problem drinking. That is not to say that a direct comparison between countries in, for example, morning drinking or blackouts can be made. As noted in the Data and Methods section there are differences in mode of collection, response rates and age ranges which cannot be controlled for in this study.

The main outcomes of the analyses on correlations between items of the AUDIT show the following:

(i) In most countries internal consistency tends to decrease or remain almost the same when the three items indicating drinking pattern are combined with the consequence items.
(ii) Frequency of drinking is an item which decreases the alpha in eight of the nine countries.
(iii) Of the consequence items, injury and concern of others lead to a decrease in alpha in, respectively, three of the eight (injury) and three of the seven (concern of others) countries which included these items.

The outcome that some items may decrease the alpha indicates that the summation of scores of the AUDIT has different

meanings in different countries. In countries where 'injury due to drinking' and/or 'frequency of drinking' or other items lower the Cronbach's alpha, the scoring on these items should not be interpreted in terms of adding to the severity of the problems measured with the other items.

Considering that there is no country in which all items contribute to the internal consistency, one could conclude that the psychometric properties of the AUDIT are not sufficient to serve as an instrument for comparable cross-national estimates of problem drinking. However, in our opinion such a conclusion would not do justice to the outcome that in most countries most items can be used in a summation score. Frequency of drinking, from a national and cross-national point of view, does not seem suitable to be included in a screening instrument used in general populations. Most likely the informative value of frequency of drinking lies in a combination of frequency with a measure of number of glasses per occasion indicating volume of consumption. In future studies it would be good to analyse whether the internal consistency of the AUDIT would increase when frequency is included in a measure of volume of consumption rather than used as a separate item.

When 'injury' and 'concern of others' also appear in other studies to show weak correlations with the other consequence items, it may be advisable not to include these items in summation scores. However, for more definitive advice as to whether or not to exclude particular consequences from an instrument measuring problem drinking, more studies are needed. The study of Kypri *et al*. (2002) shows that AUDIT items may be interpreted differently than intended when used in general population samples rather than in a more controlled treatment setting. One may hypothesize that in a cross-national context there may be more items for which there are systematic differences between countries with regard to how these items are interpreted (even when following a standard format) and that, therefore, the responses are not comparable. For European countries it has been shown that there are cultural differences between countries in the cognitive strategies used to answer standard questions about drinking (Raitasalo *et al*., 2005). Differences between countries in intercorrelations between consequences might be due to differences in how these items are understood. Studies which combine qualitative and quantitative psychometric methods (e.g. see Kypri *et al*., 2002; Raitasalo *et al*., 2005) are needed before making decisions about which items to include or exclude in an instrument allowing cross-national comparable estimates of problem drinking. The basic issue is, of course, whether on the basis of such studies it would become clear that cultural differences in interpretation of items and cultural differences in problems attributed to alcohol are that large that a standard instrument is not feasible from a cross cultural point of view. On this point we think that cultural variation might be less for items measuring immediate consequences of drinking (e.g. blackouts, morning drinking, and unable to stop once started) compared with items in which the consequence is mediated by psychological (e.g. guilt and remorse) or social reactions (e.g. concern of others) or structural and cultural differences in where drinking takes place (e.g. injury due to drinking). Especially for consequences mediated by psychological or social reactions the cultural variation between countries may be that large that there is no standard way to measure such consequences. Or, to say it differently, for instruments to be responsive to cultural differences the standard should perhaps be to allow for variation in items and wording of items.

Table 4. Prevalences (%) and gender ratios for AUDIT indicators of drinking pattern and level

	Frequency: ⩾2 per week			Quantity/ occasion: ⩾5			⩾6 drinking ⩾1 per month		
	M	F	R	M	F	R	M	F	R
Switzerland	79.4	55.6	1.4	11.6	2.5	4.6	7.6	1.2	6.3
Spain	59.0	28.0	2.1	19.6	6.4	3.1			
U.K.	63.7	44.4	1.4	7.0	1.0	7.0			
Sweden	20.3	12.5	1.6	27.1	7.7	3.5	34.3	12.5	2.7
Finland	60.6	37.2	1.6	39.4	16.9	2.0	53.5	18.4	2.9
The Netherlands	48.7	25.1	1.9	23.5	5.8	4.1	30.9	7.9	3.9
Czech Republic	62.0	30.7	2.0	60.2	24.9	2.4	26.2	8.0	3.3
Hungary	43.2	11.2	3.9	19.0	1.7	11.2	35.2	9.2	3.8
Iceland	36.8	18.8	2.0	33.1	21.3	1.6	35.3	18.0	2.0

M = male; F = female; R = gender ratio.

Table 5. Item total correlations and Cronbach's alpha coefficients of the AUDIT items

	Switzerland	Spain	UK	Sweden	Finland	The Netherlands	Czech Republic	Hungary	Iceland
Unable to stop	0.36	0.46	0.50	0.41	0.51	0.17	0.47	0.44	0.49
Normative expect	0.33	0.45	0.52	0.47	0.48	×	0.58	0.50	0.40
Morning drinking	0.32	0.48	0.34	0.34	0.47	×	0.56	0.48	0.42
Guilt/remorse	0.39	0.55	0.49	0.47	0.58	0.32	0.48	0.50	0.08
Blackout	0.39	0.51	0.55	0.54	0.61	0.39	0.59	0.46	0.53
Injured	0.12	0.20	×	0.29	0.43	0.17	0.41	0.12	0.28
Concern of others	0.25	0.24	×	0.41	0.45	0.39	0.05	–0.05	×
Frequency: ⩾2 per week audit 3–4	0.17	0.21	0.24	0.40	0.29	0.28	0.38	0.38	0.28
Quantity/occasion: ⩾5 audit 2–4	0.29	0.36	0.31	0.46	0.49	0.46	0.40	0.36	0.36
⩾6 drinking ⩾1 per month audit 2–4	0.39	×	×	0.56	0.62	0.51	0.51	0.42	0.54
Alpha Cronbach: conseq. only	0.61	0.70	0.74	0.69	0.78	0.59	0.75	0.68	0.73
Alpha Cronbach: + drink. indic.	0.58	0.62	0.69	0.71	0.81	0.63	0.77	0.68	0.76
Worst items	Injury freq.	Injury concern of others freq.	Morn. dr. freq.	Freq.	Freq.	Freq.	Concern of others	Injury concern of others freq.	Guilt/remorse freq.

Acknowledgements — This study was carried out with the cooperation of Ladislav Csemy, Czech Republic; Pia Mákelá, Finland; Zsuzsanna Elekes, Hungary; Hildigunnur Ólafsdóttir, Iceland; Antonio Vidal Infer, Spain; Karin Helmersson-Bergmark, Sweden; Gerhard Gmel, Switzerland, Moira Plant, UK. This research was conducted within the framework of a Concerted Action (QLG4-CT-2001-0196) funded by the Quality of Life and Management of Living Resources Programme of the European Commission. The data are also a part of the broader project, Gender, Alcohol, and Culture: An International Study (GENACIS). GENACIS is a collaborative international project affiliated with the Kettil Bruun Society for Social and Epidemiological Research on Alcohol and coordinated by GENACIS partners from the University of North Dakota, the University of Southern Denmark, Charité—University Medicine Berlin, Germany, the World Health Organization, and the Swiss Institute for the Prevention of Alcohol and Drug Problems. Support for other aspects of the project comes from the US National Institute on Alcohol Abuse and Alcoholism/National Institutes of Health (Grant Numbers R01AA04610 and R21AA12941, Sharon C. Wilsnack, principal investigator), the World Health Organization (Maristela Monteiro and Isidore Obot, coordinators), the German Federal Ministry of Health and Social Security, and Swiss national funds. Support for individual country surveys was provided by government agencies and other national sources. Please see the introduction of this issue for a listing of country-specific funding for these analyses. Data coordinator for the GENACIS project is Gerhard Gmel, Swiss Institute for the Prevention of Alcohol and Drug Problems, Lausanne, Switzerland.

REFERENCES

Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B. *et al*. (2001) *The Alcohol Use Disorder Identification Test; Guideline For Use in Primary Care*. WHO, Department of Mental health and substance dependence, Geneva.

Beresford, T. P., Blow, F. C., Hill, E. *et al.* (1990) Comparison of CAGE questionnaire and computer assisted laboratory profiles in screening for covert alcoholism. *Lancet* **336**, 482–485.

Cahalan, D. (1976) *Problem Drinkers*. Jossey Bass, San Fransisco.

Conigrave, K. M., Saunders, J. B. and Reznik, R. B. (1995) Predictive capacity of the AUDIT questionnaire for alcohol related harm. *Addiction* **90**, 1479–1485.

Davis, L. J., Hurt, R. D., Morse, R. M. *et al*. (1987) Discriminant analysis of the self-administered alcoholism screening test. *Alcohol, Clinical and Experimental Results* **11**, 269–273.

European Commission. (2006) *Health Determinants, Lifestyle, Alcohol*. Retrieved 14 July 2006 from http://ec.europa.eu/health/ph_determinants/life_style/alcohol_en.htm.

Fleming, J. (1996) The epidemiology of alcohol use in Australian women: findings from a national survey of women's drinking. *Addiction* **91**, 1325–1334.

Gmel, G., Heeb, J. and Rehm, J. (2001) Is frequency of drinking an indicator of problem drinking? A psychometric analysis of a modified version of the alcohol use disorder identification test in Switzerland. *Drug and Alcohol Dependence* **64**, 151–163.

Holmilla, M. (1995) Intoxication and hazardous use of alcohol: result from the 1992 Finnish Drinking Habit Study. *Addiction* **90**, 785–792.

Ivis, F. J., Adlaf, E. M. and Rehm, J. (2000) Incorporating the AUDIT into a general population telephone survey: a methodological experiment. *Drug and Alcohol Dependence* **60**, 97–104.

Knibbe, R. A., Derickx, M., Gmel, G. *et al.* (2003) *The measurement of alcohol-related problems in general population surveys in European countries*. Paper presented at the Genacis meeting, November 2003, Berlin.

Kypri, K., McGee, R., Saunders, J. B. *et al.* (2002) Interpretation of items in the AUDIT questionnaire. *Alcohol and Alcoholism* **37**, 465–467.

Mayfield, D., McLeod, G. and Hall, P. (1974) The CAGE questionnaire: validation of a new alcoholism screening instrument. *American Journal of Psychiatry* **131**, 1121–1123.

Medina-Mora, E., Carreno, S. and De la Fuente, J. R. (1998) Experience with the alcohol use disorders identification test (AUDIT) in Mexico. In *The Consequences of Alcoholism*, Vol. 14, Galanter, M. ed., pp. 383–396. Plenum Press, New York.

Raitasalo, K., Knibbe, R. and Kraus, L. (2005) Retrieval strategies and cultural differences in answering survey questions on drinking: a cross-national comparison. *Addiction Research and Theory* **13**, 359–372.

Saunders, J. B., Aasland, O. G., Amundsen, A. *et al.* (1993a) Alcohol consumption and alcohol-related problems among primary health care patients: WHO collabarative project on early detection of persons with harmful alcohol consumption I. *Addiction*, **88**, 348–362.

Saunders, J. B., Aasland, O. G., Babor, T. F. *et al.* (1993b) 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.

Selzer, M. L., Vinokur, A. and van Roooijen, L. (1975) A self-administered Short Michigan Alcoholism Screening Test (SMAST). *Journal of Studies on Alcohol* **36**, 117–126.

SPSS Inc. (2002). SPSS 11.0 for Windows. SPSS Inc., Chicago, IL.