



<http://www.diva-portal.org>

This is the published version of a paper published in *Drugs: education prevention and policy*.

Citation for the original published paper (version of record):

Bloomfield, K., Hope, A., Kraus, L. (2013)
Alcohol survey measures for Europe: A literature review.
Drugs: education prevention and policy, 20(5): 348-360
<http://dx.doi.org/10.3109/09687637.2011.642906>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-74364>

Alcohol survey measures for Europe: A literature review

Kim Bloomfield¹, Ann Hope², & Ludwig Kraus^{3,4}

¹Centre for Alcohol and Drug Research, Aarhus University, Artillerivej 90, 2300 Copenhagen S, Denmark,

²Department of Public Health and Primary Care, Trinity College Centre for Health Science, AMNCH,

Tallaght, Dublin 24, Ireland, ³IFT Institut für Therapieforschung, Parzivalstraße 25, 80804 München,

Germany, and ⁴Centre for Social Research on Alcohol and Drugs (SoRAD), Stockholm University, SE-106 91, Stockholm Sweden

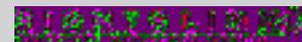
This article reports the results of a purposive review of the different alcohol survey instruments currently used in survey research. The review was conducted to support the development of a standardized comparative survey methodology for Europe. It examines various types of instruments used to measure alcohol consumption, risky drinking, alcohol abuse and dependence, social consequences and third-party harm. The review of literature provides several insights for a European comparative survey. The beverage-specific quantity–frequency measure is recommended for use across countries. A reference period of 1 year for alcohol consumption is considered important if one is to link associated problems with alcohol intake. With regard to risky drinking, objective measures based on the quantity of approximately 60–70 g of ethanol per drinking occasion are preferable to subjective measures of drunkenness. In choosing an instrument for measuring abuse and dependence, the key issue is to decide whether the instrument is to serve as a screening or diagnostic tool. In the case of screening, the Alcohol Use Disorders Identification Test or the Rapid Alcohol Problem Screen-4 appears more appropriate. But if one desires to approximate a diagnostic category, then the Composite International Diagnostic Interview or another operationalization of ICD-10/DSM-IV criteria would be the better choice. Due to a lack of validated scales for social consequences and third-party harm, no recommendations are justified.

INTRODUCTION

Alcohol is a major health determinant in the European Union (EU). Presently, it is estimated that 53 million EU adults (15%) do not drink alcohol at all and some 58 million (16%) are heavy drinkers, of whom some 23 million (6.5%) are considered dependent on alcohol (Anderson & Baumberg, 2006). In the recent Eurobarometer alcohol survey, almost 133 million EU citizens reported risky drinking (≥ 50 g of alcohol per occasion) at least once a month, representing one in three of the adult population (Taylor Nelson Sofres PLC Opinion and Social, 2010). However, the prevalence of harmful drinking patterns depends to a large extent on the questions asked, definitions and methodologies used (Anderson & Baumberg, 2006). Currently, across the EU there is no standardized comparative survey methodology available for measuring alcohol intake in general as well as for heavy drinking, episodic heavy drinking (binge drinking), and alcohol abuse and dependence.

Some work has been done through the European Community Health Indicators for Monitoring project (ECHIM) that is conducting, among others, an assessment of the comparability of all health indicators endorsed by the European Commission. ECHIM has to date proposed total alcohol consumption and hazardous alcohol consumption as risk factor indicators. Important efforts are being made within the European Health Interview Survey (EHIS) to elaborate standardized questions on alcohol consumption and to apply them in its large health survey. The EU Committee on alcohol data, indicators and definitions recommends three key indicators for alcohol consumption and harm: volume of consumption, pattern of consumption and alcohol-attributable health harm. The volume of alcohol consumption is defined as total consumption per

Copyright Clearance Center
Not for Sale
Authorized for personal use only
displayed on a single copy for personal use.



capita (≥ 15 years), pattern of consumption is defined as harmful drinking (≥ 60 g per occasion) at least monthly and alcohol-attributable health harm is measured as alcohol-attributable years of life lost (premature mortality) (Committee on Data Collection IaD, 2009).

In its 2007 work plan, the European Commission called for the development of standardized comparative surveys on drinking patterns. A major aim of the EU project 'Standardized Measurement of Alcohol-Related Troubles' (SMART) has been to develop a comprehensive comparative survey instrument on alcohol use, patterns of drinking and alcohol problems for use in the EU. The SMART project co-funded by the European Commission Directorate General for Health and Consumers (DG-SANCO) was set up in 2008 with experts involved in alcohol survey methodology in 10 Member States. This article reports the results of a purposive review of various alcohol survey instruments currently used in survey research. The review was conducted to support the development of a standardized comparative survey methodology for Europe. It examines the various types of instruments used to measure alcohol consumption including measures of risky drinking. The review also focuses on three main areas of alcohol-related harm: abuse/dependence, negative social consequences and third-party harm.

METHODS

The approach used for this review was purposive or realist, meaning that the search strategy was adapted to the project's particular purpose, i.e. to review relevant literature to help guide a research project on alcohol measures for application within the EU. Two major reviews on alcohol measures were used as a point of departure (Dawson, 2003; Gmel & Rehm, 2004). In these, previous literature had been thoroughly addressed and discussed. From them, we have highlighted particularly relevant subjects and studies and have added more recent literature that has appeared since these reviews were conducted. With respect to the literature on alcohol-related problems, we have consulted experts in order to grasp the universe of problem measurement. To this, we have used our own knowledge of the field in order to begin our purposive search.

ALCOHOL CONSUMPTION MEASURES

Frequency and quantity

The main approaches to measuring alcohol intake in survey research can be categorized in the following way: quantity–frequency (QF) measures; graduated-frequency (GF) measures; and short-term recall measures. The simplest measure of alcohol consumption is one based purely on frequency. This has been the case in medical epidemiology studies (Rehm, 1998) and in basic population health surveys. However, such a

measure does not allow the calculation of volume of alcohol consumed. The QF measure allows for calculation of volume since quantity consumed is recorded along with the frequency of drinking. Although this measure has a long history (e.g. Straus & Bacon, 1953), it can be considered one of the most universal and practical instruments in alcohol survey research. However, it has been criticized to measure modal values of frequency and quantity instead of averages, thus not yielding a true average total volume (Dawson, 2003). Modifications or improvements have been added to it, such as a beverage-specific version which is asked of the three main beverage types. Also, an additional question about binge or episodic heavy drinking is usually added to the basic measure to accommodate wider variability in drinking patterns. The QF measure may be asked for varying periods: usually anywhere from a week to a year, depending upon the nature of the research question and the main drinking pattern of the population under study.

The second main measure, the GF, takes a different approach which implicitly has the goal of measuring volume. Numbers of drinks consumed on an occasion are grouped into graduated categories. In its most modern version, the GF asks a respondent the maximum number of drinks he/she has had in a specified period. After it has been established which category of drinks is the highest for that respondent, he/she is then asked how often that has occurred. Then, the respondent is asked progressively in groups of fewer drinks how often he/she has drunk such amounts on an occasion. In such a way, the research attempts to cover the entire 'universe' of one's drinking volume by starting with the maximum amount and working down by asking about smaller amounts. This approach has been in use since the early 1960s and is employed mainly in North America. In its modern version, the GF is generally not beverage-specific. The period of questioning is usually the past year. One of its strengths is that it can more easily identify occasions of heavy consumption. In 2003, a task force of the Council of the US National Institute on Alcohol Abuse and Alcoholism (NIAAA) identified a minimum set of questions mainly deriving from the QF and GF approaches that is designed to capture both consumption and drinking pattern. The set offers options for three (containing the basic QF and a question on consuming five or more drinks on an occasion), four, five and six items, and can be useful when space and time are of a premium (NIAAA, 2003).

The final main approach used in alcohol survey research is short-term recall measures, such as the *weekly recall* (WR). In this approach, respondents are basically asked to recall all alcohol they have consumed in a recent short period, such as the previous week. Due to the short period, it is assumed that respondents will be able to correctly recall all their consumption during that period. This approach is rather easy to administer and for respondents to understand.



The main drawback is the short measurement interval which is especially disadvantageous in capturing the full patterns of infrequent drinkers who may, for example, have just had a week of either no consumption or of unusually heavy consumption. Patterns and volume based on such an interval would misclassify these drinkers. Recently, other short-term recall methods have been more extensively assessed: for instance, the 'Yesterday' method (Stockwell, Zhao, Chikritzhs, & Greenfield, 2008). In the detailed version of this measure, respondents are asked about what they consumed the day before, including beverage types, sizes and alcohol content. This approach also minimizes under-reporting and offers high coverage of official national per capita consumption estimates and information about risks associated with different beverage types. Unfortunately, short-term recall methods are weak with regard to examining individual drinking patterns because they fail to recognize changes in drinking behaviour over time. While they accurately assess consumption of regular drinkers, drinking patterns of occasional drinkers are not captured well.

An extensive literature exists on comparative assessments of the above-mentioned instruments. Key studies that are relevant and informative for this study are listed in Table I. First to be considered are the comparisons between the generic QF measure and its extended version. The basic finding of both Kühlnhorn and Leifman (1993) and Williams, Proudfit, Quinn, and Campbell (1994) was that when more questions are asked, more volume is reported. In the study of Kühlnhorn and Leifman (1993), a QF measure that divided the days of the week into Monday–Thursday, Friday, Saturday and Sunday yielded higher volume than a measure of a 'normal week's consumption'. Williams et al. (1994) compared a generic QF measure with a beverage-specific quantity–frequency (BSQF) as well as a BSQF with specific drink sizes. This resulted in increasingly higher volume with increasing elaboration of the generic QF. Additional studies illustrate the main principle that with increasing the number of questions on consumption, the recorded volume of alcohol will likewise increase (Dawson, 1998; Knibbe & Bloomfield, 2001; Rehm, 1998). Moreover, a classical study by Lemmens, Tan, and Knibbe (1992) was able to demonstrate that in comparison with the diary method (as the 'gold standard') the source of under-reporting among the other tested measures of 7-day recall and QF lay in the frequency dimension (rather than the quantity dimension). This source of bias is stronger among those who drink infrequently.

Several studies have been conducted comparing the QF to the GF. An early study found no major differences between the two measures (Hilton, 1989), in which a BSQF and general GF were compared to a 10-week prospective diary. Midanik (1994) found that the GF captured a 38% higher volume than the generic QF. Later studies conducted outside the USA have found complications with the GF. Although the

measure yields higher volume than the generic QF, its use often results in over-counting of drinking frequencies, including substantial proportion of respondents whose estimated annual frequency surpassed 365 days (Gmel et al., 2006; Graham et al., 2004; Poikilainen et al., 2002). Thus, it is not clear whether the resulting higher volume is a valid result or artefact. In an international comparison across 10 countries, Gmel et al. (2006) found problems with implementation of the GF in that in several countries respondents reacted to only one quantity category. The BSQF appeared to perform better by yielding higher volumes in most cases and was not prone to as much double-counting of drinking frequencies as the GF. In summary, the GF appears to work better with lighter drinkers and among those with more cognitive skills since the median of drinking is being asked as well as the task of dividing up a total of drinking days across levels of volume. If a wide variety of societies is to be compared with the same instrument, Gmel et al. (2006) find the BSQF to function better. Finally, several studies have examined the QF in relation to retrospective diaries or short-recall methods. Here, the results are the most consistent of all comparisons in that diaries yield larger volumes, especially for those categorized as lighter drinkers (Redman et al., 1987; Shakeshaft et al., 1999; Werch, 1989; Webb et al., 1990). However, it should be noted that such an instrument has the limitation of a short-recall period and cannot be applied for larger intervals as can the summary measure of the QF.

Very closely related to choice of alcohol intake measure is the consideration of the reference period for that intake (Dawson, 2003). As one can see, some of the measures have an inherent period 'built in' to the instrument. For example, retrospective diary measures cover, by necessity, rather short periods in which drinking occasions can be accurately recalled. Thus, with these measures, the trade-off is between a high degree of recall accuracy at the expense of a short reference period. A short reference period has the disadvantage of not adequately capturing the drinking pattern and volume of especially light and infrequent drinkers (Dawson, 2003). On the other hand, longer periods require respondents to summarize or average their drinking quantity and frequency. In particular, in the QF approach, underestimation of both measures is very likely, and high volume consumption is excluded by definition. To some extent, the same is true for the GF, however, respondents are also actually asked to remember occasions on which they drank the most alcohol in a given period. Periods of 1 month to a year are also common for the GF. Another consideration is that a longer reference period is necessary if one is to also attempt to associate problems with alcohol intake. Enough time needs to have occurred beforehand in order to argue that a risk exposure of alcohol is associated with problem consequences (Dawson, 2003).



Table I. Overview of key studies comparing alcohol intake measures.

Comparison	Study	Country	Results/conclusions
Generic QF vs. extended QF	Kühlhorn and Leifman (1993)	Sweden	Compared weekday-specific QF with generic QF. Day-specific yielded higher volumes
	Williams et al. (1994)	USA	Compared QF and BSQF, the BSQF gave higher volume estimates
GF vs. QF	Hilton (1989)	USA	No significant differences found between GF and BSQF
	Midanik (1994)	USA	Higher volume for GF than generic QF
	Poikilainen, Podkletnova, and Alho (2002)	Finland	Higher volume for GF than generic QF but over-counted annual frequency
	Graham, Demers, Rehm, and Gmel (2004)	Canada	GF yielded higher volume but over-counted frequency
	Gmel, Graham, Kuendig, and Kuntsche (2006)	International	BSQF performed better in international comparison. GF gave higher frequencies. Implementation problems
QF vs. retrospective diaries	Redman, Sanson-Fisher, Wilkinson, Fahey, and Gibberd (1987)	Australia	Diaries yield more at lower volumes
	Werch (1989)	USA	Diaries yield higher frequencies and volume for 7-day recall
	Webb, Redman, Sanson-Fisher, and Gibberd (1990)	Australia	Diaries yield more at lower volumes for 7-day recall
	Shakeshaft, Bowman, and Sanson-Fisher (1999)	Australia	Diaries yield higher volume for 7-day recall
Reviews	Rehm (1998)	Canada	'The more specific the questions asked about alcohol consumption, the more volume results'. Does not prefer GF or QF
	Fuenekes, van't Veer, van Staveren, and Kok (1999)	The Netherlands	Favour a BSQF if diary not possible
	Dawson (2003)	USA	Considers GF and QF as main measures
	Gmel and Rehm (2004)	Switzerland	Comprehensive review: conclude that 'the selection of best measure will depend on purpose of the study'
	Greenfield and Kerr (2008)	USA	Make no recommendation on type of measure but stated that 'GF is an efficient measure in countries where individuals drink smaller quantities per occasion'

Risky drinking

Terms to describe risky drinking include episodic heavy drinking, binge drinking, risky single occasion drinking and extreme drinking. Because average daily intake and consequently average consumption may not adequately reflect risks associated with certain outcomes, a measure of more intensive, concentrated consumption taking place within a short period has become recognized as a critical measure of alcohol drinking pattern (Dawson, Li, & Grant, 2008). Such a measure has come to be defined generally as that which can increase blood alcohol concentration to a level of intoxication within an occasion. Thus, this has been seen to be a dose of approximately 60–70 g of ethanol. The number of standard drinks necessary to reach this level depends on the ethanol content of a standard drink which may vary across countries. In the USA, where a standard drink contains 13.6 g ethanol, the equivalent

in drinks is roughly four to five drinks. In some countries, sex-specific levels are set; for instance, for some US researchers, five or more drinks are used for men and four or more for women (e.g. Wechsler, Dowdall, Davenport, & Rimm, 1995). Table II provides an overview of defined quantities constituting risky drinking in surveys collected in the Gender, Alcohol and Culture: An International Study project (GENACIS).

In order to empirically support such quantities, research, at least in the USA, has attempted to evaluate the association between drinking levels and levels of risk (Dawson, 2000; Dawson, Grant, & Li, 2005; Dawson et al., 2008). Dawson et al. (2005) constructed risk curves to determine the relationship of consumption exceeding five drinks for men and four drinks for women with the incidence of alcohol abuse and/or dependence. A clear, almost linear relationship



Table II. Overview of risk drinking definitions from the GENACIS project (volume per drinking occasion).

Country	Risk drinking measure definition: number of standard drinks (ethanol in grams)	
	<i>Standard drinks/glasses</i>	<i>Grams (approximately)</i>
Switzerland	8+	80 for men and women
Germany	5+	70 for men and women
Israel	5+	60
Mexico	5+	65
Sweden	6+	72
Finland	6+/5+ from GF measure	60
The Netherlands	6+	60
Czech Republic	5+glasses	90
Hungary	Sum of frequencies drinking 3–5 and 6+ drinks	60
Brazil	5+	60
Iceland	5+	65
Denmark	6+	72
Sri Lanka	5+	60
Nigeria	5+	60
Kazakhstan	5+	60
Argentina	5+	60
Canada	5+	68
USA	5+	60–70
Uganda	5+	60
Japan	Sum of frequencies for 6–9 and 10+ units	72
Costa Rica	5+	60
India	5+	50
Australia	6+	60
Ireland	6+	60
Uruguay	5+	60
Belize	5+	50
Peru	5+	60
Australia	6+	60
New Zealand	6+	60

Source: Gender, Alcohol and Culture: An International Study (2009).

between number of days exceeding the limits and the incidence of dependence with abuse could be seen when graphed. Thus, at least based on American research, heavy drinking cut-off points have a justification based on empirical evidence.

An alternative approach to assess risky drinking is to ask respondents for the experience of drunkenness which has been discussed as a subjective operationalization of binge drinking (Gmel, Rehm, & Kuntsche, 2003). Compared to objective measures of volume or frequency of episodic heavy drinking, drunkenness may contain information on behavioural and physical effects that consumption measures may not capture, i.e. it might be self-adjusted for individual differences in

body water percentage and alcohol metabolism (Kerr, Greenfield, & Midanik, 2006). In empirical studies, the subjective measure has proved to be a better predictor of social consequences, alcohol dependence symptoms and alcohol-related harm than five or more drinks and feeling the effects (Midanik, 1999). Despite these potential advantages, drunkenness seems to be unsuitable when it comes to cross-national studies due to its strong proneness to individual, environmental and cultural influences (Greenfield & Kerr, 2008). In a qualitative study, Cameron et al. (2000) were able to demonstrate different cultural connotations of the word drunkenness with English, Dutch and Swedish people mainly associating behavioural terms, whereas Scots and Greeks used psychological definitions.

MEASURES OF ALCOHOL-RELATED HARM AND NEGATIVE CONSEQUENCES

One of the main problems of the measurement of alcohol-related harm and negative consequences concerns the fact that most questions ask if the incidence has occurred because of the presence of alcohol. It has been stated that only in alcohol epidemiology, the measurement of the risk factor (alcohol) is already implicitly associated with the outcome (Rehm & Gmel, 1999; Room, 2000). This is problematic for proper epidemiological analyses where the risk factor and outcome should be measured independently of each other as two separate entities. Most recently, Gmel, Kuntsche, Wicki, and Labhart (2010) found that self-attributed alcohol problems were under-reported if compared to non-attributed problems using alcohol-attributable fractions. Despite the debate that has occurred over several years, no consensus to delete the alcohol-attribution from questions on alcohol-related consequences has been reached. Although the topic is not handled further here, readers should be aware of the issue with respect to potential analyses and assignation of attribution or association.

Alcohol abuse and dependence

Measures of alcohol abuse and dependence mainly have origins in psychiatric, epidemiologic and public health research. Some work has existed in sociological survey research, but in the past 20 years this territory has become increasingly occupied by larger health authorities (e.g. World Health Organization (WHO), NIAAA) to systematically develop valid and somewhat universal instruments. The development of diagnostic criteria and corresponding instruments to operationalize these criteria has a complicated history that is too extensive to be related here. The reader is referred to Hasin (2003) in order to gain some background.

Three main measures presently tend to dominate population survey research on alcohol: the *Composite International Diagnostic Interview Substance Abuse Module* (CIDI-SAM; Üstün et al., 1997), the *Alcohol Use Disorder and Associated Disabilities Interview*



Schedule (AUDADIS; Grant, Harford, Dawson, Chou, & Pickering, 1995), and the *Alcohol Use Disorders Identification Test* (AUDIT; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993). All were developed roughly around the same time; however, only the CIDI and the AUDIT have been meant for international, comparative use. For whatever reason, the AUDADIS appears to have been used almost solely in the USA.

Shorter instruments exist such as the *CAGE* (Ewing, 1984), the *Alcohol, Smoking and Substance Involvement Screening Test* (ASSIST; Newcombe, Humeniuk, & Ali, 2005), the *Fast Alcohol Screening Test* (FAST; Hodgson, Alwyn, John, Thom, & Smith, 2002), the *Brief Michigan Alcoholism Screening Test* (B-MAST; Pokorny, Miller, & Kaplan, 1972) and the *TWEAK* (Russell et al., 1994). Although brief, these focus on special subsets of the general population. The CAGE was developed for use in clinical settings focusing on dependence only; the ASSIST is suitable for use in primary care settings identifying problematic and risky alcohol use, whereas the FAST was developed for emergency departments and other medical settings and screens for hazardous drinking and dependence. The B-MAST detects chronic alcoholics in clinical settings and the TWEAK was originally designed to screen pregnant women for harmful drinking habits but has also been used in in- and outpatient clinical settings and in the general population. The use of some of these instruments appears to largely remain in North America and thus may reflect their cultural specificity. None has been designed for use in general population surveys and only few have been validated in this context. Another brief screening tool, the *Rapid Alcohol Problem Screen* (RAPS4; Cherpitel, 2000), was initially developed to quickly identify problem drinkers among emergency room patients but has subsequently been used in general population survey research (Cherpitel, 2002). This tool has been applied to international research and has shown promising results.

This review will thus focus on the CIDI and the AUDIT as potential survey measurement instruments for detecting alcohol abuse and dependence in a European context. It also examines the RAPS as a brief screening tool that could find wider use in Europe. From the outset, however, it is important to note that the three instruments have different goals in their measurement capabilities: the CIDI sets out to replicate and operationalize DSM-IV and ICD-10 criteria for alcohol abuse and dependence so that epidemiologists in particular may identify cases on an international basis within large general population mental health surveys (Üstün et al., 1997). The background of the AUDIT lies more in previous screening instruments and not in reflecting DSM and ICD criteria per se. It was derived from testing 150 different items on primary health care patients in six countries and is meant to screen for 'persons with hazardous or harmful alcohol

consumption before dependence and serious harm have occurred' (Saunders et al., 1993, p. 791). Further, the RAPS was intended to be an even briefer tool aimed at indicating dependence for screening in the emergency room setting. Despite their varying intentions, all three instruments have been used in survey and outpatient research to detect problem and/or dependent drinkers. Table III, adapted from Hasin (2003), gives an overview of the dimensions covered by each instrument.

A review of the item coverage for the instruments makes clear that the CIDI comprehensively covers both abuse/harmful use and dependence symptoms as determined by DSM-IV and ICD-10 criteria, whereas the AUDIT can be seen as focusing on the beginning signs of trouble. The RAPS4 (and earlier original RAPS) shows most similarity to items in the older CAGE screener and with some of the AUDIT items.

The validity and reliability of the CIDI (as well as *Schedules for Clinical Assessment in Neuropsychiatry* (SCAN) and AUDADIS) was tested in 10 countries (Üstün et al., 1997). Kappa statistics were calculated to assess test-retest reliability as well as concurrent validity. The CIDI alcohol component demonstrated acceptable reliability for both DSM-IV and ICD-10 criteria as well as concurrent agreement with the SCAN and AUDADIS. Hasin (2003) has reported on later assessments as acceptable. However, the main obstacle in using the CIDI is its length and complexity.

A group in Germany developed the *Munich Composite International Diagnostic Interview* (M-CIDI) (Lachner et al., 1998). The questions assessing the diagnostic criteria are identical to those in the original CIDI but in other parts, the M-CIDI shows several improvements compared to the CIDI. For example, the screening section covering QF measures was extended by adding circumstances of first use, effects and problems associated with the initiation of alcohol use. Information about onset and recent substance use were separated into a series of questions. Moreover, more visual aids were introduced in order to simplify the identification of different beverages and to help in the presentation of complex questions. This instrument also demonstrated acceptable kappa values for reliability and validity. A novel feature is that the instrument was tested on a community sample and not among patients. In the last decade, CIDI has been applied to general population samples in many countries throughout the world.

Since its development and publication in 1993 (Saunders et al., 1993), the AUDIT has undergone extensive testing and review (e.g. Allen, Litten, Fertig, & Barbor, 1997; Berner, Kriston, Bentele, & Härtner, 2007; Reinert & Allen, 2002). Interestingly, over the years the use of the AUDIT or parts of it have extended beyond primary health care settings to general population surveys (e.g. Bloomfield, Wicki, Gustafsson, Mäkelä, & Room, 2010; Fleming, Barry, & McDonald, 1991; Gmel, Heeb, & Rehm, 2001; Homila, 1995; Ivis, Adlaf, & Rehm, 2000; Knibbe, Derickx,



Table III. Overview of diagnostic criteria with CIDI, AUDIT and RAPS4 instruments.

<i>DSM-IV dependence criteria (American Psychiatric Association [DSM-IV], 1994)</i>	<i>ICD-10 dependence criteria (WHO, [ICD-10], 1992)</i>	CIDI	AUDIT	RAPS4
Withdrawal	Withdrawal	✓		
Tolerance	Tolerance	✓	✓	
Drinking more or longer than intended	Drinking larger amounts or over a longer period than intended/persistent desire to cut down or control alcohol use	✓		
Persistent desire or unsuccessful efforts to cut down		✓	✓	
Important activities given up or reduced	Important alternative pleasures or interests reduced or given up/great deal of time spent obtain or take alcohol or getting over its effects	✓		
Great deal of time spent on drinking or getting over its effects		✓		
Continued drinking despite knowledge of serious problems	Persistent use despite evidence of harmful consequences	✓		
	Strong desire/compulsion to take the substance	✓		
<i>DSM-IV substance abuse criteria</i>				
Inability to fulfil roles		✓	✓	✓
Hazardous/dangerous use		✓	✓	
Recurrent legal problems		✓		
Continued use despite social interpersonal problems		✓		
<i>ICD-10 harmful use criteria</i>				
Pattern of use causing damage to health. The damage may be physical (e.g. liver cirrhosis) or mental (e.g. episodes of depressive disorder secondary to heavy consumption of alcohol)				
<i>Other criteria</i>				
Remorse (CAGE)			✓	✓
Blackout			✓	✓
Morning drinking (CAGE)				✓
Been told to cut down (quasi-CAGE, TWEAK)			✓	

Kuntsche, Grittner, & Bloomfield, 2006; Medina-Mora, Carreño, & de la Fuente, 1998; Selin, 2006; Wicki, Gustafsson, Mäkelä, & Gmel, 2009). Its major advantage over the CIDI is its brevity of 10 items.

Allen et al. (1997) and Reinert and Allen (2002, 2007) have regularly reviewed the literature on the AUDIT's performance. These reviews maintain that the instrument has sound psychometric properties, acceptable reliability and validity, is easy to use, relatively free of cultural bias and is obtainable without copyright fees. As of 2007, Reinert and Allen (2007) report that a larger evidence base exists for the English

version of the questionnaire; further psychometric work is urged on other language versions and gender-specific cut-offs should be considered. Additionally, a recent meta-analysis was conducted on 19 studies. Berner et al. (2007) report that results varied due to the variety of patient settings. Although sensitivities and specificities were all in the acceptable range, they urge its use be restricted to primary care, inpatient and elderly patient settings.

Selin (2006) undertook a comprehensive analysis of the instrument's validity with respect to its potential to measure various components of problematic drinking: high-volume consumption, social problems, health



problems and dependence (as tested against ICD-10 criteria). From her review of previous research it is unclear what the definitive factor structure of the instrument is. Some researchers have found a two-factor structure for consumption and problems (Kelly & Donovan, 2001), whereas others have found a three-factor solution involving consumption, harmful use and dependence (Rist, Glöckner-Rist, & Demmel, 2009). Still another group has found four factors: one for each consumption item and another for problems (Gmel et al., 2001). Selin's work did not involve finding factors but testing the instrument against the four-mentioned components. In all, she determined that the AUDIT performed well against all four criteria with a cut-off of 8+ for the whole instrument (i.e. AUDIT-10).

Curiously, in reviewing the previous work neither Selin nor the other authors described any factor as representing dependence as such. Indeed, the AUDIT was not meant to be a diagnostic instrument for dependence (Saunders et al., 1993). Its mistaken use in measuring dependence (vs. hazardous or harmful alcohol consumption) as well as its growing application in general population surveys are deviations from its original purpose. The AUDIT represents an instrument of great popularity and may have been promoted and used by many because of its international roots. However, researchers should remain mindful of the instrument's limitations (Midanik, Greenfield, & Bond, 2007).

In one of the few studies to examine the AUDIT in general population surveys, Knibbe et al. (2006) reported on the reliability of individual items in nine European countries with a special focus on gender. The conclusions were that among the consumption items, frequency of drinking decreased Cronbach's alpha correlation in almost all countries. Of the consequence items, it appeared that some questions (especially injury and concern of others) also reduced the internal consistency leading the authors to conclude that these have varying meaning across the study countries.

Finally, the RAPS4 (and its antecedent, RAPS) has shown good psychometric properties (Cherpitel et al., 2005) and has been successfully used in a series of emergency room studies in various countries such as Poland (Cherpitel, Moskalewicz, Swaitkiewicz, Ye, & Bond, 2009), Argentina (Cremonte & Cherpitel, 2008) and Mexico (Cherpitel & Borges, 2000). It also performed well in the US general population survey for alcohol dependency but less so for alcohol abuse (Cherpitel, 2002). However, when QF questions (drinking five or more drinks on at least one occasion during the last year and drinking as often as once a month during the last year) were added to the RAPS4, sensitivity for alcohol abuse was improved. The instrument has been the measurement basis in large international studies encompassing countries from North and South America, Europe, Asia and Africa (e.g. Cherpitel et al., 2005). In its international

application, the RAPS4 has demonstrated higher sensitivity and specificity in relation to a measure of tolerance (proxy for dependence) in those countries with more problematic drinking cultures (i.e. where alcohol is not well integrated into the culture) (Cherpitel et al., 2005).

Social consequences

There is a tradition to view separately those alcohol-related problems which concern abuse and dependence (formerly alcoholism), and those which concern social problems that the drinker experiences due to his or her consumption (e.g. problems with spouse/relatives/friends, with job, with police, with finances, with aggression). These have been labelled in various ways; for example, as 'intrinsic' and 'extrinsic' (Bloomfield et al., 2010), or 'internal' and 'external' (Plant, Miller, Thornton, Plant, & Bloomfield, 2000), and 'preoccupation with alcohol' and 'troubles due to drinking' (Mulford & Miller, 1960), 'dependence' and 'consequences' (Hilton, 1991) and 'personal consequences' and 'social consequences' (Rootman & Moser, 1985). All dichotomies reflect the intra- and inter-personal dimensions of the problems. However, at least two studies have shown that when psychometrically analysing such items for evidence of their dimensionality, researchers tend to find only one main component for all items (Gmel, Rehm, Room, & Greenfield, 2000; Wicki et al., 2009). Thus from a statistical standpoint, all alcohol-related problems appear to cohere to one main dimension even though researchers would like to separate and categorize them on theoretical or conceptual grounds.

In the case of this review, the former approach of categorization has been used to organize this section and thus its topic now concerns what can be considered those external or extrinsic problems which deal with difficulties in the social realm. As Room (2000) has remarked, there is little consensus on standard instruments for measuring the social consequences of one's drinking; now more commonly called 'social harm'. Room defines social harm as 'perceived mis-performance or failure to perform in major social roles – as a family member, as a worker, as a friend or neighbour, or in terms of public demeanour' (Room, 2000, p. 94). Further, he remarks that over the last half-century, the number of items measuring these areas has varied yet the number of areas has remained stable. In general, these areas for which questions have been asked over the last five decades include problems with: spouse, relatives, neighbours and friends, job, police, aggression and belligerence, finances. In addition to the areas, there are levels of 'intensity' at which these problems can be measured. Room (2000, p. 106) observes the following levels of external or extrinsic problems, starting from the most intense:

- (1) concrete actions reported by others in response to respondent's drinking,



Table IV. Summary of life areas for social consequences.

	Level 1	Level 2	Level 3	Level 4
Life area	Concrete action	Verbal responses, attempts to control	Global attribution	R's behaviour while drinking seen as problem
Spouse	My wife left me because of my drinking	My wife threatened to leave me because of my drinking	My drinking has had a harmful effect on my marriage	I almost had sex with another woman at my wife's birthday party
Relatives	A relative does not speak to me anymore because of my drinking	A relative has indicated I should cut down on my drinking	My drinking has had a harmful affect on my relatives	I offended my mother in-law after having a couple of drinks
Friends and neighbours	A friendship ended because of my drinking	A friend indicated I should cut down	My drinking has had a harmful effect on a friendship	I have said harsh or cruel things to someone/ friend while drinking
Job	I have lost a job because of my drinking	People at work have indicated I should cut down	My drinking has hurt my chances for promotion or better jobs	I had to take a day of sick leave after drinking
Police	I have been arrested because of my drinking	A policeman questioned or warned me about my drinking	I have had trouble with the law because of my drinking	By chance, they failed to stop me for drunk driving being occupied with another driver
Aggression/ belligerence	I have gotten into a fight while drinking; I have felt aggressive or cross while drinking	A bartender warned me to call the police next time I misbehave	I have had trouble with violence due to my drinking	I almost hit my baby after a few drinks
Finances	I had to sell my jewellery to pay my drinking debts	Co-owner of our business told me to reduce drinking	My drinking has had a harmful effect on my finances	I drank away money intended for new TV set

Note: Adapted from Room (2000).

- (2) indications of verbal responses to or attempts to control respondent's drinking,
- (3) global attribution by the respondent of harm in a life area,
- (4) items describing the respondent's behaviour while drinking as problematic.

According to Room's categorizations, Table IV illustrates with selected items how questions in the various life areas at different levels of intensity manifest themselves. Note that some of the questions are hypothetical and are not actually survey items that have been empirically or psychometrically tested.

Third-party harm

A relatively new area of alcohol-related problems in alcohol survey research is third-party harm (Room, 2000; Rossow & Hauge, 2004). This area of problems, measured at the individual level, has extremely little in the way of systematic review or documentation. The earliest item-list on third-party harm in alcohol survey research appears to come from a Canadian study in 1989 (Eliany, Giesbrecht, Nelson, Wellman, & Wortley, 1992). Around the same time, a Nordic alcohol survey contained items on social harm from others (Mäkelä et al., 1999). From these two surveys, items are summarized in Table V. Rossow and Hauge (2004) have analysed the Nordic survey items for

Norway to examine the extent of harm by others. They report that the more severe types of harm are experienced by few (*ca.* 3–5%) but the less severe types (e.g. kept awake at night) have a considerable prevalence; i.e. *ca.* 20% report that they have been kept awake at least once in the past 12 months. They also conclude that those who experience the most third-party harm also are heavy drinkers, and that women experience more of this kind of harm than do men (Rossow & Hauge, 2004).

The problem items considered so far are those which would be included in a general battery for a national alcohol survey. More extensive lists that concentrate on a specific dimension of third-party problems such those experienced by family members (e.g. Orford, Templeton, Velleman, & Copello, 2005) or as intimate partner violence (Bergmark, Graham, & Nordvik, 2005) are few and little research, as mentioned, is to be found in the literature.

CONCLUSIONS

The review of literature provides several insights for the SMART project on measurement issues. A basic finding for alcohol consumption measures is that when more questions are asked, more volume is reported. The BSQF approach is recommended when a wide



Table V. Listing of existing batteries of third-party harm from alcohol surveys.

	Problem area
<i>Canadian survey (Eliany et al., 1992)</i>	
Have you ever	
Been insulted or humiliated by someone who had been drinking?	Insult
Had serious agreements or quarrels as a result of someone else's drinking?	Aggression
Had friendships break up as a result of someone else's drinking?	Friendship
Had family problems or marriage difficulties due to some else's drinking?	Family/marriage
Been a passenger with a driver who had too much to drink?	Driving
Been in a motor vehicle accident because of someone else's drinking?	Driving
Had your property vandalized by someone who had been drinking?	Vandalism
Been pushed, hit or assaulted by someone who had been drinking?	Violence
Been disturbed by loud parties or the behaviour of people drinking?	Disturbance
Had financial trouble because of someone else's drinking?	Finances
<i>Nordic survey (Mäkelä et al., 1999)</i>	
Has it ever happened that	
You have been harassed or bothered by intoxicated people on the street or in some other public place?	Insult public
You have been harassed or bothered by intoxicated people at a party or some other private setting?	Insult private
An intoxicated person has harmed you physically?	Violence
An intoxicated person has ruined your clothes or other belongings?	Vandalism
You have been called names or otherwise insulted by intoxicated people?	Insult
You have been afraid of intoxicated people you encountered on the street?	Fear
You have been kept awake at night by drunken noises?	Disturbance

variety of societies are to be compared using the same instrument. A reference period of 1 year for alcohol consumption is considered important if one is to link associated problems with alcohol intake.

The objective measure of risky drinking based on the quantity of approximately 60–70 g of ethanol per drinking occasion was found preferable to the subjective measure of drunkenness. Even though there are some advantages to drunkenness measures, e.g. potential self-adjustment for individual differences, there is strong evidence for cultural variation in the perception of drunkenness based on the same amount of alcohol (Müller, Piontek, Kraus, & Pabst, 2011). Thus, such measures are not suitable for international comparative studies. Accurate measures of risky drinking based on volume and pattern of consumption are suited for international comparisons and as indicators of exposure for risk assessments of alcohol-related mortality and morbidity (Rehm et al., 2003, 2009).

The key review finding in choosing an instrument for measuring abuse/dependence is to decide whether the chosen instrument is to serve as a screening tool or as a diagnostic tool. In the case of screening for dependence, the AUDIT and the RAPS4 appear to be appropriate instruments due to their brevity. But if one desires to approximate a diagnostic category, then CIDI or another operationalization of ICD-10 or DSM-IV criteria would be the better choice. As a consequence of the planned revision of the diagnostic systems (i.e. DSM-V), the operationalizations of alcohol use disorder will change (O'Brian, 2011). This will require adaptations of screening instruments and further

validation studies. As a positive side effect, this may offer opportunities for sound psychometric analyses on an international level.

Measures of social consequences and third-party harm basically cover the same underlying concept of social harm but apply different perspectives. In the first case, harms are executed by the (drinking) respondent to others, whereas in the second case, the harm is done by (drinking) others to the respondent. Presently, there is little consensus on standard instruments. Due to this lack of validated scales, survey research is mainly based on single items. Based on these considerations, no specific recommendations are justified. However, surveys may be used for the development and evaluation of psychometrically sound instruments.

ACKNOWLEDGEMENTS

Funding for Kim Bloomfield and Ludwig Kraus within the project 'Standardizing Measurement of Alcohol Related Troubles (SMART)' was provided by the Commission of the European Communities in the framework of the Public Health Programme (Grant Agreement No. 2007308). Ludwig Kraus was also co-funded by the German Federal Ministry of Health (Grant II45-2508DSM304). Ann Hope was supported by the Department of Health and Children and by the Health Service Executive in Ireland.



Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

REFERENCES

- Allen, J., Litten, R., Fertig, J., & Barbor, T. (1997). A review of research on the Alcohol Use Disorders Identification Test (AUDIT). *Alcoholism: Clinical and Experimental Research*, *21*, 613–619.
- American Psychiatric Association (1994). *DSM-IV: Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.
- Anderson, P., & Baumberg, B. (2006). *Alcohol in Europe: A public health perspective*. London: Institute of Alcohol Studies.
- Bergmark, K.H., Graham, K., & Nordvik, M. (2005). A comparison of alcohol-related aggression in six European countries. In K. Bloomfield, A. Allamani, F. Beck, K.H. Bergmark, L. Csémy, I. Eisenbach-Stangl, . . . , M. Romero Mendoza (Eds.), *Gender, culture and alcohol problems: A multi-national study* (pp. 103–115). Berlin: Charité University Medicine.
- Berner, M., Kriston, L., Bentele, M., & Härter, M. (2007). The alcohol use disorders identification test for detecting at-risk drinking: A systematic review and meta-analysis. *Journal on Studies of Alcohol and Drugs*, *68*, 461–473.
- Bloomfield, K., Wicki, M., Gustafsson, N.-K., Mäkelä, P., & Room, R. (2010). Changes in alcohol-related problems after alcohol policy changes in Denmark, Finland and Sweden. *Journal on Studies of Alcohol and Drugs*, *71*, 32–40.
- Cameron, D., Thomas, M., Madden, S., Thornton, C., Bergmark, A., Garretsen, H., & Terzidou, M. (2000). Intoxicated across Europe: In search of meaning. *Addiction Research*, *8*, 233–242.
- Cherpitel, C. (2000). A brief screening instrument for problem drinking in the emergency room: The RAPS4. *Journal of Studies on Alcohol*, *61*, 447–449.
- Cherpitel, C. (2002). Screening for alcohol problems in the US general population: Comparison of the CAGE, RAPS4 and RAPS4-QF by gender, ethnicity and service utilization. *Alcoholism-Clinical and Experimental Research*, *26*, 1686–1691.
- Cherpitel, C.J., & Borges, G. (2000). Screening instruments for alcohol problems: A comparison of cut points between Mexican American and Mexican patients in the emergency room. *Substance Use and Misuse*, *35*, 1419–1430.
- Cherpitel, C., Moskalewicz, J., Swiatkiewicz, G., Ye, Y., & Bond, J. (2009). Screening, Brief Intervention and Referral to Treatment (SBIRT) in a polish emergency department: Three-month outcomes for a randomized, controlled clinical trial. *Journal of Studies on Alcohol and Drugs*, *70*, 982–990.
- Cherpitel, C., Ye, U., Bond, J., Borges, G., Cremonte, M., Marias, S., . . . , Swiatkiewicz, M. (2005). Cross-national performance of the RAPS4/RAPS-QF for tolerance and heavy drinking: Data from 13 countries. *Journal of Studies on Alcohol*, *66*, 428–432.
- Committee on Data Collection IaD (2009). *Summary report*. Brussels: European Commission. Retrieved from http://ec.europa.eu/health/indicators/events/alcohol_committee_mtg_2_2009-09_minutes_final.pdf
- Cremonte, M., & Cherpitel, C.J. (2008). Performance of screening instruments for alcohol use disorders in emergency department patients in Argentina. *Substance Use and Misuse*, *43*, 125–138.
- Dawson, D. (1998). Volume of ethanol consumption: Effects of different approaches to measurement. *Journal of Studies on Alcohol*, *59*, 191–197.
- Dawson, D. (2000). US low-risk drinking guidelines: An examination of four alternatives. *Alcoholism: Clinical and Experimental Research*, *24*, 1820–1829.
- Dawson, D. (2003). Methodological issues in measuring alcohol use. *Alcohol Research and Health*, *27*, 18–29.
- Dawson, D., Grant, B., & Li, T. (2005). Quantifying the risks associated with exceeding recommended drinking limits. *Alcoholism: Clinical and Experimental Research*, *29*, 902–908.
- Dawson, D., Li, T., & Grant, B. (2008). A prospective study of risk drinking: At risk for what? *Drug and Alcohol Dependence*, *95*, 62–72.
- Eliany, M., Giesbrecht, N., Nelson, M., Wellman, B., & Wortley, S. (1992). *Alcohol and other drug use by Canadians: A national alcohol and other drug survey*. Ottawa: Health Canada.
- Ewing, J. (1984). Detecting alcoholism: The CAGE questionnaire. *Journal of the American Medical Association*, *252*, 1905–1907.
- Fleming, M., Barry, K., & McDonald, R. (1991). The Alcohol Use Disorders Identification Test (AUDIT) in a college sample. *International Journal of the Addictions*, *26*, 1173–1185.
- Fuenekes, G.I., van't Veer, P., van Staveren, W.A., & Kok, F.J. (1999). Alcohol intake assessment: The sober facts. *American Journal of Epidemiology*, *150*, 105–112.
- Gender Alcohol and Culture: An International Study (2009). *Recommendations for the use of drinking indicators*. Retrieved from <http://www.genacis.org/index.php10>
- Gmel, G., Graham, K., Kuendig, H., & Kuntsche, S. (2006). Measuring alcohol consumption - should the 'graduated frequency' approach become the norm in survey research? *Addiction*, *101*, 16–30.
- Gmel, G., Heeb, J., & Rehm, J. (2001). Is frequency of drinking an indicator of problem drinking? A psychometric analysis of a modified version of the alcohol use disorders identification test in Switzerland. *Drug and Alcohol Dependence*, *64*, 151–161.
- Gmel, G., Kuntsche, E., Wicki, M., & Labhart, F. (2010). Measuring alcohol-related consequences in school surveys: Alcohol-attributable consequences or consequences with students' alcohol attribution. *American Journal of Epidemiology*, *171*, 93–104.
- Gmel, G., & Rehm, J. (2004). Measuring alcohol consumption. *Contemporary Drug Problems*, *31*, 467–540.
- Gmel, G., Rehm, J., & Kuntsche, E. (2003). Binge drinking in Europe: Definitions, epidemiology, and consequences. *Sucht*, *49*, 105–116.
- Gmel, G., Rehm, J., Room, R., & Greenfield, T. (2000). Dimensions of alcohol-related social and health consequences in survey research. *Journal of Substance Abuse*, *12*, 113–138.
- Graham, K., Demers, A., Rehm, J., & Gmel, G. (2004). Problems with the graduated frequency approach to measuring alcohol consumption: Results from a pilot study in Toronto, Canada. *Alcohol and Alcoholism*, *39*, 455–462.
- Grant, B., Harford, T., Dawson, D., Chou, P., & Pickering, R. (1995). The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): Reliability of alcohol and



- drug modules in a general population sample. *Drug and Alcohol Dependence*, 39, 37–44.
- Greenfield, T.K., & Kerr, W.C. (2008). Alcohol measurement methodology in epidemiology: Recent advances and opportunities. *Addiction*, 103, 1082–1099.
- Hasin, D. (2003). Classification of alcohol use disorders. *Alcohol Research and Health*, 27, 5–17.
- Hilton, M. (1989). A comparison of a prospective diary and two summary recall techniques for recording alcohol consumption. *British Journal of Addiction*, 84, 1085–1092.
- Hilton, M. (1991). Demographic characteristics and the frequency of heavy drinking as predictors of self-reported drinking problems. In W. Clark & M. Hilton (Eds.), *Alcohol in America: Drinking practices and problems* (pp. 194–212). Albany, NY: State University of New York Press.
- Hodgson, R., Alwyn, T., John, B., Thom, B., & Smith, A. (2002). The fast alcohol screening test. *Alcohol and Alcoholism*, 37, 61–66.
- Homila, M. (1995). *Intoxication and hazardous use of alcohol: Results from the 1992 Finnish drinking habits study*. Helsinki: Social Research Institute on Alcohol Studies.
- Ivis, F., Adlaf, E., & Rehm, J. (2000). Incorporating the AUDIT into a general population telephone survey: A methodological experiment. *Drug and Alcohol Dependence*, 60, 97–104.
- Kelly, T., & Donovan, J. (2001). Confirmatory factor analysis of the Alcohol Use Disorder Identification Test (AUDIT) among adolescents treated in emergency departments. *Journal of Studies on Alcohol*, 62, 838–842.
- Kerr, W.C., Greenfield, T.K., & Midanik, L.T. (2006). How many drinks does it take you to feel drunk? Trends and predictors for subjective drunkenness. *Addiction*, 101, 1428–1437.
- Knibbe, R., & Bloomfield, K. (2001). Alcohol consumption estimates in surveys in Europe: Comparability and sensitivity for gender differences. *Substance Abuse*, 22, 23–38.
- Knibbe, R., Derickx, M., Kuntsche, S., Grittner, U., & Bloomfield, K. (2006). A comparison of the Alcohol Use Disorder Identification Test (AUDIT) in general population surveys in nine European countries. *Alcohol and Alcoholism*, 41, i19–i25.
- Kühlhorn, E., & Leifman, H. (1993). Alcohol surveys with high and low coverage rate: A comparative analysis of survey strategies in the alcohol field. *Journal of Studies on Alcohol*, 54, 542–554.
- Lachner, G., Wittchen, H.-U., Perkonig, A., Holly, A., Schuster, P., Wunderlich, U., . . . Pfister, H. (1998). Structure content and reliability of the Munich-Composite International Diagnostic Interview (M-CIDI) substance use sections. *European Addiction Research*, 4, 28–41.
- Lemmens, P., Tan, E., & Knibbe, R. (1992). Measuring quantity and frequency of drinking in a general population survey: A comparison of five indices. *Journal of Studies on Alcohol*, 53, 476–486.
- Mäkelä, P., Fonager, K., Hibell, B., Nordlund, S., Sabroe, S., & Simpura, J. (1999). *Drinking habits in the Nordic countries* (Report No. 2/99). Oslo: National Institute for Alcohol and Drug Research.
- Medina-Mora, M.E., Carreño, S., & de la Fuente, J. (1998). Experience with the Alcohol Use Disorders Identification Test (AUDIT) in Mexico. In M. Galanter, & L.A. Kaskutas (Eds.), *Recent developments in alcoholism* (Vol. 14, pp. 383–397). New York: Plenum Press.
- Midanik, L.T. (1994). Comparing usual quantity/frequency and graduated frequency scales to assess yearly alcohol consumption: Results from the 1990 US National Alcohol Survey. *Addiction*, 89, 407–412.
- Midanik, L.T. (1999). Drunkenness, feeling the effects and 5+ measures. *Addiction*, 94, 887–897.
- Midanik, L.T., Greenfield, T.K., & Bond, J. (2007). Addiction sciences and its psychometrics: The measurement of alcohol-related problems. *Addiction*, 102, 1701–1710.
- Mulford, H., & Miller, D. (1960). Drinking in Iowa: IV. Preoccupation with alcohol and definitions of alcohol heavy drinking and trouble due to drinking. *Quarterly Studies on Alcohol*, 60, 203–208.
- Müller, S., Piontek, D., Kraus, L., & Pabst, A. (2011). The relationship between alcohol consumption and perceived drunkenness: A multilevel cross-national comparison in samples of adolescents. *Alcohol and Alcoholism*, 46, 399–406.
- Newcombe, D.A.L., Humenik, R.E., & Ali, R. (2005). Validation of the World Health Organization Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): Report of results from the Australian site. *Drug and Alcohol Review*, 24, 217–226.
- National Institute on Alcohol Abuse and Alcoholism (NIAAA) (2003). Task force on recommended alcohol questions – National Council on Alcohol Abuse and Alcoholism recommended sets of alcohol consumption questions - October 15–16, 2003. Bethesda, MD: Author. Retrieved from <http://www.niaaa.nih.gov/Resources/ResearchResources/TaskForce.htm>
- O'Brian, C. (2011). Addiction and dependence in DSM-V. *Addiction*, 106, 866–867.
- Orford, J., Templeton, L., Velleman, R., & Copello, A. (2005). Family members of relatives with alcohol, drug and gambling problems: A set of standardized questionnaires for assessing stress, coping and strain. *Addiction*, 100, 1611–1624.
- Plant, M., Miller, P., Thornton, C., Plant, M., & Bloomfield, K. (2000). Life stage, alcohol consumption patterns, alcohol-related consequences and gender. *Substance Abuse*, 21, 265–281.
- Poikolainen, K., Podkletnova, I., & Alho, H. (2002). Accuracy of quantity-frequency and graduated frequency questionnaires in measuring alcohol intake: Comparison with daily diary and commonly used laboratory markers. *Alcohol and Alcoholism*, 37, 573–576.
- Pokorny, A., Miller, B., & Kaplan, H. (1972). The brief MAST: A shortened version of the Michigan Alcoholism Screening Test. *American Journal of Psychiatry*, 129, 342–345.
- Redman, S., Sanson-Fisher, R., Wilkinson, C., Fahey, P., & Gibberd, R. (1987). Agreement between two measures of alcohol consumption. *Journal of Studies on Alcohol*, 48, 104–108.
- Rehm, J. (1998). Measuring quantity, frequency and volume of drinking. *Alcoholism: Clinical and Experimental Research*, 22 (Suppl. S2), 4S–14S.
- Rehm, J., & Gmel, G. (1999). Patterns of alcohol consumption and social consequences. Results from an 8-year follow-up study in Switzerland. *Addiction*, 94, 899–912.
- Rehm, J., Mathers, C., Popova, S., Thavorncharoensap, M., Teerawattananon, Y., & Patra, J. (2009). Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. *Lancet*, 373, 2223–2233.
- Rehm, J., Room, R., Graham, K., Monteiro, M., Gmel, G., & Sempos, C.T. (2003). The relationship of average volume of



- alcohol consumption and patterns of drinking to burden of disease: An overview. *Addiction*, 98, 1209–1228.
- Reinert, D., & Allen, J. (2002). The Alcohol Use Disorders Identification Test (AUDIT): A review of recent research. *Alcoholism: Clinical and Experimental Research*, 26, 272–279.
- Reinert, D., & Allen, J. (2007). The alcohol use disorders identification test: An update of research findings. *Alcoholism: Clinical and Experimental Research*, 31, 185–199.
- Rist, F., Glöckner-Rist, A., & Demmel, R. (2009). The Alcohol Use Disorders Identification Test revisited: Establishing its structure using nonlinear factor analysis and identifying subgroups of respondents using latent class factor analysis. *Drug and Alcohol Dependence*, 100, 71–82.
- Room, R. (2000). Concepts and items in measuring social harm from drinking. *Journal of Substance Abuse*, 12, 93–111.
- Rootman, I., & Moser, J. (1985). *Community response to alcohol-related problems*. Washington, DC: US Government Printing Office.
- Rossow, I., & Hauge, R. (2004). Who pays for the drinking? Characteristics of the extent and distribution of social harms from others' drinking. *Addiction*, 99, 1094–1102.
- Russell, M., Martier, S., Sokol, R., Mudar, P., Bottoms, S., Jacobson, S., & Jacobson, J. (1994). Screening for pregnancy risk-drinking. *Alcoholism: Clinical and Experimental Research*, 18, 1156–1161.
- Saunders, J., Aasland, O., Babor, T., de la Fuente, J., & 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.
- Selin, K.H. (2006). Alcohol Use Disorder Identification Test (AUDIT): What does it screen? Performance of the AUDIT against four different criteria in a Swedish population sample. *Substance Use and Misuse*, 41, 1881–1899.
- Shakeshaft, A., Bowman, J., & Sanson-Fisher, R. (1999). A comparison of two retrospective measures of weekly alcohol consumption: Diary and quantity/frequency index. *Alcohol and Alcoholism*, 34, 636–645.
- Stockwell, T., Zhao, J., Chikritzhs, T., & Greenfield, T. (2008). What did you drink yesterday? Public health relevance of a recent recall method used in the 2004 Australian National Drug Strategy Household Survey. *Addiction*, 103, 919–928.
- Straus, R., & Bacon, S. (1953). *Drinking in college*. New Haven, CT: Yale University Press.
- Taylor Nelson Sofres PLC Opinion & Social (2010). *EU citizens' attitudes towards alcohol. Special Eurobarometer* (Report No. 72.3). Brussels: European Commission.
- Üstün, B., Compton, W., Mager, D., Babor, T., Baiyewu, O., & Chatterji, S. (1997). WHO study on the reliability and validity of the alcohol and drug use disorder instruments: Overview and methods. *Drug and Alcohol Dependence*, 47, 161–169.
- Webb, G., Redman, S., Sanson-Fisher, R., & Gibberd, R. (1990). Comparison of a quantity-frequency method and a diary method of measuring alcohol consumption. *Journal of Studies on Alcohol*, 51, 271–277.
- Wechsler, H., Dowdall, G., Davenport, A., & Rimm, E. (1995). A gender-specific measure of binge drinking among college students. *American Journal of Public Health*, 85, 982–985.
- Werch, C. (1989). Quantity-frequency and diary measures of alcohol consumption for elderly drinkers. *International Journal of the Addictions*, 24, 859–865.
- Wicki, M., Gustafsson, N.-K., Mäkelä, P., & Gmel, G. (2009). Dimensionality of drinking consequences - cross-cultural comparability and stability over time. *Addiction Research and Theory*, 17, 2–16.
- Williams, G., Proudfit, A., Quinn, E., & Campbell, K. (1994). Variations in quantity-frequency measures of alcohol consumption from a general population survey. *Addiction*, 89, 413–420.
- WHO (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva: Author.

