THE FACTOR STRUCTURE AND CONCURRENT VALIDITY OF THE ALCOHOL USE DISORDER IDENTIFICATION TEST BASED ON A NATIONALLY REPRESENTATIVE UK SAMPLE

MARK SHEVLIN* and GILLIAN W. SMITH

Psychology Research Institute, University of Ulster, Northland Road, Londonderry, Northern Ireland BT48 7JL, UK

(*Received* 4 *January* 2007; *first review notified* 18 *February* 2007; *in revised form* 23 *April* 2007; *accepted* 24 *April* 2007; *advance access publication* 28 *July* 2007)

Abstract — Aims: To assess the psychometric structure and construct validity of the alcohol use disorder identification test (AUDIT) in a Great Britain population sample. **Methods:** A stratified multi-stage random sample of 7849 participants completed the AUDIT as part of a computer assisted interview. Confirmatory factor analyses were conducted testing one to three factor models. The factors in these models were correlated with demographic variables and scores relating to perceived wellbeing, verbal IQ, and neurotic and psychosis symptoms to assess construct validity of the factor solutions. **Results:** A two factor solution was deemed to appropriately fit the data, measuring alcohol consumption and alcohol related problems. Correlations between the two factors on demographic, wellbeing, neurosis and psychosis symptomology were significantly different. **Conclusions:** The two factor solution suggests an advantage to investigating factor specific cut off scores for both consumption and alcohol related problems given their difference in predictive validity on both health and demographic variables.

The Alcohol Use Disorder Identification Test (AUDIT) was designed as a brief screening instrument to detect hazardous drinking behaviours in the past year (Reinert and Allen, 2002). It is one of the most commonly used screening instruments and aims to detect problem drinking rather than established alcoholism (Allen *et al*., 1997).

The original AUDIT represented three conceptual domains; consumption (items 1–3), dependence (items 4–6) and alcohol related consequences (items 7–10) suggesting three underlying dimensions (Saunders and Aasland, 1987). Despite this, the scoring of the scale involves summing all items to provide a possible range of score from 0 to 40, with a score of 8 or more being suggested as an indicator of hazardous drinking behaviour. This scoring method implies a unidimensional (one factor) scale structure with equal weight given to each item in computing the final score. Subsequent evidence regarding the construct validity of the AUDIT, in particular the number of dimensions that are being measured, has been equivocal. Support for one, two and three factor models have been based on both exploratory and confirmatory factor analytic methods (CFA).

A one dimensional structure of the questionnaire was supported by Skipsey *et al*. (1997) based on a principal components analysis (PCA) of hazardous drinking in drug dependent patients ($N = 82$). El-Bassel *et al*. (1998) found a similar structure in a population of women offenders in a New York City jail facility ($N = 400$), also using PCA. Carey *et al*. (2003) also drew this conclusion in an investigation of the factor structure of the AUDIT in the language most suitable for their sample of psychiatric patients in India ($N = 671$) using both PCA and CFA on the structure predicted by the exploratory procedure. In all these studies the one factor represented hazardous drinking.

There has been greater research support for a two-factor model. Medina-Mora *et al*. (1998) explored the factor structure using a translated version of the AUDIT in a non-clinical male Mexican population ($N = 2050$). Using exploratory factor analysis with varimax rotation, they reported a two factor solution, representing patterns of intake (items 1–3) and alcohol problems (items 4–9). Item 10 loaded on both factors. O'Hare and Sherrer (1999) also used a PCA with varimax rotation which supported a two factor model explaining 60% of the variance. They used a sample of undergraduates from the University of Rhode Island ($N = 312$). The two factors were similar in structure to the Medina-Mora model, although item 8 loaded on both factors and item 10 loaded highly on factor two.

Chung *et al*. (2002) performed a confirmatory factor analysis testing one and three factor structures ($N = 173$), on a version of the AUDIT which modified items 3, 6 and 8 for their population of adolescents presenting to an emergency department. They performed a subsequent PCA with varimax rotation and concluded the two factor structure to be superior, with items 1–3 loading on factor one (labelled 'consumption') and 4–10 on factor two (labelled 'problems'). Kelly and Donovan (2001) used confirmatory factor analyses to test one, two and three factor models based on scores from 103 adolescents. A correlated two factor solution similar to that of Chung *et al*. (2002) was found to be the best fitting model, however, issues regarding model identification led to the exclusion of item 9 from the analysis. Karno *et al*. (2000) performed a confirmatory factor analysis based on a predominantly male (86%) sample of mental health clinic outpatients ($N = 197$). One and three factor models were tested and neither model provided adequate fit. Subsequent analyses using PCA, principal axis factoring (PAF) and maximum likelihood (ML) exploratory factor analyses indicated that a two-factor model represented the best fitting model (although in the PCA and PAF items 4 and 5 loaded on both factors, and in the ML model item 9 did not significantly load on either factor). Broadly speaking, this suggests that the two

*Author to whom correspondence should be addressed at: Psychology Research Institute, University of Ulster, Northland Road, Londonderry, BT48 7JL. Tel: +44 (0) 28 71 375619; E-mail: m.shevlin@ulster.ac.uk

factor solution of items 1–3 as consumption measures and the remaining items measuring problems exists for this population, however, some disagreement is evident within the results for items 4, 5 and 9. Maisto *et al*. (2000) in a PCA with varimax rotation (and supported by confirmatory factor analyses) on transformed data from primary care patients in Pennsylvania ($N = 7035$) found two factors; consumption (items 1–3) and dependence/consequences (items 4–9), with item 10 loading on both factors. Using samples from college students ($N = 465$) and court referred outpatients in community mental health and substance use treatment centre ($N = 135$), Shields *et al*. (2004) also performed confirmatory factor analyses on the one, two and three factor models, and concluded superior fit for a correlated two factor solution in both populations on the basis of parsimony and chi-squared difference testing. Like many of the other papers drawing this conclusion, these two factors measured consumption (questions 1–3) and dependence/consequences (items 4–10).

Three papers note the factor structure of the AUDIT in a randomly selected general population sample. Conclusions by Bergman and Kallmen (2002) in a general Swedish population sample ($N = 997$) found a two factor solution (labelled 'hazardous consumption' – items 1–3 and 'alcohol related problems' items 4–10) using both CFA and EFA. Lima *et al*. (2005) in Brazil ($N = 166$) also concluded a superior fit for the two factor model using CFA and PCA methods on a translated version of the AUDIT which accounted for existing Portuguese translations of the instrument. These two factors were concluded to measure both consumption (questions 1–3) and alcohol related problems (questions 4–10). A four factor solution was proposed by Gmel *et al*. (2001) using confirmatory factor analyses on a modified version of the AUDIT. Modifications included a rewording of questions 1 and 2 to represent quantity frequency per occasion, with question 3 asking how often the individual had 8 drinks or more. This paper concluded the best structure as items 1 to 3 loading on one latent variable each with the remainder on a fourth latent variable. This would support the two factor model in terms of items 1, 2 and 3 being distinct from the rest, although, it is unclear as to how the latent variables can be statistically identified when measured with a single observed variable.

Although there has been consistent support for a two factor structure of the AUDIT there has been some support for a three factor model. Maisto *et al*. (2000) and Shields *et al*. (2004) both reported acceptable fit for both the two and three factor solutions. In both cases the two factor solution was preferred based on both parsimony and the reported high factor correlations between factors two and three (measuring dependence and consequences), which suggested factor redundancy. Whilst this provides an argument for selecting the two factor solution, the adequacy of the three factor model should not be discounted until after the unique predictive validity of each of the factors in both structures have been tested. Finally, it must be noted that studies using a modified or translated version of the AUDIT should be cautiously interpreted in relation to those using an original version.

The aim of this paper was to test the dimensionality, and assess the concurrent validity, of the AUDIT based on a large nationally representative sample of British participants. This study will provide information on the psychometric functioning of the AUDIT as a screening instrument within community surveys. The validity of the factors underlying the AUDIT was assessed by means of correlations with demographic variables and scores relating to perceived wellbeing, verbal IQ, neurotic and psychosis symptoms. On the basis of the prevalence estimates reported by Singleton *et al*. (2001) it was predicted that any factors derived from the AUDIT would be negatively associated with age and sex (indicating significantly higher scores for males). Furthermore it was predicted that AUDIT factors would positively associated with income (MacDonald and Shields, 2001; van Ours, 2004) and negatively with IQ (Windle and Blane, 1989; Mortensen *et al*., 2006). In term of the psychological and health variables it was predicted that there would be negative correlations with mental and physical health, and positive correlations with psychotic and neurotic symptoms (Jané-Llopis and Matytsina, 2006). This study analysed population data, rather than focusing on problem alcohol use. Therefore it is expected that the magnitude of the effects will be small.

METHOD

Participants and data

The analyses were conducted on data from the 'Psychiatric Morbidity of Adults living in Private Households, 2000' survey (Singleton *et al*., 2001). This research focused on adults aged 16 to 74 years, who were recruited using a stratified multi-stage random probability sampling strategy. Interviews were successfully conducted with 8580 adults. All were resident in either Scotland, England or Wales. After listwise deletion of missing data the total effective sample size was $N = 7849$. The mean age of the sample was 45 years old (SD = 15.43), 54.1% were female and 93% were of white ethnic origin. The interview was conducted using computer assisted interviewing. Full details of the survey methods can be found in Singleton *et al*. (2001).

MEASURES

The following measures were used in this study:

AUDIT (Saunders and Aasland, 1987; Babor et al., 1992)
This is a 10-item scale referring to alcohol use in the past 12 months with a score ranging from 0 to 40. Internal reliability estimates for the scale over 18 studies had a median Cronbach's alpha of above 0.80 and estimates of test-retest reliability ranging from 0.64 to 0.92 over three studies (Reinert and Allen, 2002).

Clinical Interview Schedule-Revised (CIS-R: Lewis and Pelosi, 1990; Lewis *et al*., 1992). Neurotic symptoms were measured using this scale and based on experiences in the week prior to interview. The CIS-R measures 14 neurotic symptoms according to the ICD-10 criteria and is summed to give a total CIS-R score. Higher scores are indicative of more neurotic symptoms being reported.

The National Adult Reading Test (NART: Nelson and O'Connell, 1978; Nelson and Willison, 1991). This is a test

of the ability to read and pronounce correctly 50 words, reflecting the extent of individual's intellectual development in adulthood. Using algorithms recommended in Section 2 of the NART test manual, scores on the NART can be converted into WAIS-R Verbal IQ scores (Nelson and Willison, 1991).

Psychosis Screening Questionnaire (PSQ; Bebbington and Nayani, 1995). The PSQ was used to assess psychotic symptoms within the past year. The PSQ has five probe questions (plus secondary questions) enquiring about mania, thought insertion, paranoia, strange experiences and hallucinations. It has been found to have sensitivity 92% and specificity of 95%. High scores indicate greater endorsement of psychosis-like symptoms.

SF-12 Health Survey (Ware *et al.*, 1996). The SF-12 was designed as a brief version of the SF-36 which measures two components, physical and mental health. In general population surveys, it has been found that the two components are also the two factors in the SF-12 (Jenkinson and Layte, 1997).

ANALYSIS

A series of confirmatory factor models were specified and estimated using maximum likelihood parameter estimation. Analyses were conducted using LISREL 8.70 (Joreskog and Sorbom, 2004a). Covariance and asymptotic weight matrices were calculated using PRELIS 2.7 (Joreskog and Sorbom, 2004b). Using an asymptotic weight matrix allows for weaker assumptions in the distribution of the observed variables and by presenting a more accurate estimate of the population matrix, improving both the fit of the model and the test statistics (Satorra, 1992; Curran *et al.*, 1996).

As recommended by Hoyle and Panter (1995) goodness of fit for each of the proposed models was determined using the Satorra-Bentler scaled chi-square (S-Bχ^2), the incremental fit index (IFI; Bollen, 1989), and the comparative fit index (CFI; Bentler, 1990). A chi-square which is not significant and values in excess of 0.95 for both the IFI and CFI indicate acceptable model fit. The root-mean-square error of approximation (RMSEA; Steiger, 1990) with 90% confidence intervals were also reported, with acceptable fit for the RMSEA value less than 0.05, and values up to 0.08 indicating reasonable errors of approximation in the population (Joreskog and Sorbom, 1993). Hu and Bentler (1999) also recommend the use of the standardized root-mean-square residual (SRMR; Joreskog and Sorbom, 1981). It is considered that values of less than 0.08 indicate acceptable model fit (Hu and Bentler, 1998). Comparative fit of the models were assessed using the expected cross-validation index (ECVI; Browne and Cudeck, 1989), which is used for the purpose of comparing models with the smallest value representative of superior fit.

Three models were specified and estimated. Model 1, the unidimensional model, was implied by the scoring procedure and supported by Skipsey *et al.* (1997). Here all items were specified to load on a single factor. Model 2, the two factor model, specifies the factors of 'alcohol consumption' (items 1–3) and 'alcohol related problems' (4–10). Model 3 specified three factors, 'alcohol consumption' (items 1–3), 'dependence' (4–6) and 'related problems' (7–10). For all

Table 1. Fit indices for the alternative models of the Alcohol Use Disorder Identification Test (AUDIT)

Item	Model 1	Model 2	Model 3
S-Bχ^2	1456	488	401
df	35	34	33
P	0.000	0.000	0.000
RMSEA	0.072	0.041	0.038
90% CI	0.069–0.075	0.038–0.045	0.035–0.042
ECVI	0.60	0.18	0.14
IFI	0.96	0.99	0.99
CFI	0.96	0.99	0.99
SRMR	0.074	0.048	0.041

S-Bχ^2, Satorra-Bentler Chi-Square; RMSEA, Root mean square of approximation; ECVI, Expected cross-validation index; IFI, Incremental fit indices; CFI, comparative fit indices; SRMR, Standardized root mean square residual.

models the error covariances were fixed to zero, and all the factors were free to covary. The predictive validity was assessed by correlations between the summed scores for the factors derived from the three alternative factor models and health and demographic variables.

RESULTS

The mean AUDIT score for the entire sample was 5.68 (SD = 4.53). The mean score for men (M = 7.01; SD = 5.02) was higher than that for women (M = 4.55; SD = 3.71), this difference was statistically significant, t(1,6514.70) = 24.29, *P* <0.01. Hazardous drinking (measured as AUDIT score of 8 or more) was found in 38.8% males and 18.6% females (26.4% overall).

Confirmatory factor analysis
Table 1 reports the fit indices for the three proposed model structures.

For Model 1 the S-Bχ^2 was large relative to the degrees of freedom, although caution is exercised when interpreting the chi-square as a measure of model fit given a large sample size due to the increased power of the test (Tanaka, 1987). All other fit indices indicated acceptable fit. Models 2 and 3 also were also adequate descriptions of the sample data and provided better fit than Model 1 in terms of all the fit indices. The ECVI also indicated that Models 2 and 3 were better than Model 1 by having lower values. However, the difference between Models 2 and 3 in terms of fit was marginal. The IFI and CFI are higher, and the RMSEA and SRMR are lower for Model 3, however this model is less parsimonious than Model 2.

The factor loadings for models 2 and 3 are presented in Table 2. For both models the factor loadings are all positive, high, and statistically significant. For Model 2 the standardized loadings ranged from 0.26 to 0.92 and the factors correlate moderately at 0.54 (*P* < 0.05). For Model 3 the standardized loadings ranged from 0.26 to 0.91. The consumption and dependence factors correlated at 0.47 (*P* <

Table 2. Standardized factor loadings (standard errors) for two factor model of the AUDIT

| | Model 2 | | Model 3 | | |
	Alcohol consumption	Alcohol related problems	Alcohol consumption	Dependence	Alcohol related consequences
Item					
1. How often do you have a drink containing alcohol?	0.46 (0.01)		0.46 (0.01)		
2. How many drinks containing alcohol do you have on a typical day when you are drinking?	0.68 (0.02)		0.68 (0.02)		
3. How often do you have six or more drinks on one occasion?	0.92 (0.01)		0.91 (0.01)		
4. How often during the last year have you found that you were not able to stop drinking when you started?		0.70 (0.02)		0.72 (0.02)	
5. How often during the last year have you failed to do what was normally expected of you because of drinking?		0.71 (0.02)		0.74 (0.02)	
6. How often during the last year have you needed a first drink in the morning to get yourself going after a heavy drinking session?		0.51 (0.02)		0.53 (0.02)	
7. How often during the last year have you had a feeling of guilt or remorse after drinking?		0.65 (0.02)			0.64 (0.02)
8. How often during the last year have you been unable to remember what happened the night before because of your drinking?		0.67 (0.02)			0.69 (0.02)
9. Have you or someone else been injured because of your drinking?		0.26 (0.02)			0.26 (0.02)
10. Has a relative, friend, doctor or health worker been concerned about your drinking or suggested you cut down?		0.44 (0.02)			0.44 (0.02)
Factor Correlations Factor one	1.00		1.00		
Factor two	0.55 (0.01)	1.00	0.47 (0.01)	1.00	
Factor three			0.64 (0.01)	0.94 (0.02)	1.00

Note. All factor loadings and factor correlations statistically significant ($P < 0.05$) and standard errors in parenthesis.

0.05), the consumption and alcohol related problems factors at 0.64 ($P < 0.05$), and the dependence and alcohol related problems factors at 0.94 ($P < 0.05$). Although the correlation between the dependence and alcohol related problems factor in Model 3 is quite high, neither model can be rejected given the fit statistics in Table 1. The choice of best model should be made on the basis of parsimony and evidence of unique predictive validity of the factors. That is, the three-factor model would have to provide additional useful information compared to the two-factor model to be considered as an acceptable factor structure.

Reliability and variance explained

Reliability analyses were conducted on the three proposed model solutions. The single factor solution had a Cronbach's alpha 0.78, with the two factor solution 0.69 and 0.66 for factors one and two respectively. The three class solution had reliability of 0.69, 0.65 and 0.58 respectively for each of the three factors.

Variance explained by the single factor in the one factor solution only accounts for 30.11% of the variance in the data. Looking at the two factor solution, factor one explains 50.03% of the variance in items 1 to 3 and factor 2 explains 34.21% of the variance in items 4 to 10. Finally, within the three

factor solution, the three factors account for 58.17%, 13.61% and 13.15% respectively of the variance in the model.

Predictive validity

The unique predictive validity of each of the factors in Model 2 and 3 was assessed in terms of their correlations with health and demographic variables. The correlations for the one factor model demonstrate that a multidimensional structure offers more information than a unidimensional model. The correlations and significance values can be found in Table 3.

The aggregate score of all the items, representing the unidimensional structure (Model 1), correlates significantly with all variables except the physical component of the SF-12. The correlations of the two factors from Model 2 indicates that the factors are conceptually distinct as there are differentially associated with the criterion variables. The correlations indicate a different association between the physical and mental component of the SF-12, with consumption having a low but significant positive correlation with physical health and dependence/consequences having a low but significant negative correlation with mental health. Furthermore, PSQ total score correlates only with dependence and consequences with a significant difference between factors. All other health and demographic variables correlates significantly with factors

Table 3. Correlations between AUDIT factor scores from Models 1–3 on health measures and demographic criteria

		Health measures				Demographic criteria			
	Factor	SF-12 Physical	SF-12 Mental	Total PSQ	Total CIS-R	Age	Sex[a]	Income	Verbal IQ
Model 1	Hazardous drinking	0.02	−0.10**	0.06**	0.16**	−0.26**	−0.22**	0.08**	−0.05**
Model 2	Alcohol consumption	0.11**	0.01	0.00	0.03*	−0.33**	−0.24**	0.17**	−0.07**
	Alcohol related problems	0.01	−0.09**	0.05**	0.13**	−0.21**	−0.20**	0.04**	−0.05**
Model 3	Alcohol consumption	0.11**	0.01	0.00	0.04*	−0.33**	−0.24**	0.17**	−0.07**
	Dependence	−0.02*	−0.13**	0.07**	0.19**	−0.19**	−0.19**	0.02	−0.05**
	Alcohol related consequences	0.01	−0.11**	0.06**	0.17**	−0.25**	−0.20**	0.05**	−0.05**

Note:
* $P < 0.05$,
** $P < 0.01$;
[a] Point-biserial correlation.

one and two and are significantly different from each other except verbal IQ.

For Model 3, the correlations between all the criterion variables and the dependence and the alcohol related consequences factors are very similar. This would suggest that these factors are conceptually very similar and that no further explanatory power comes from having an additional factor. The factors are essentially interchangeable.

CONCLUSION

Three alternative factor models were proposed and tested based on data from a national survey of Great Britain. Whilst there was a slight statistical advantage to the three factor solution in terms of fit indices, the additional factor (over a two factor model) did not offer additional unique predictive validity. The high correlation between factors two and three of the three factor model resulted in almost identical correlations with all health and demographic criterion variables. Maisto *et al*. (2000) drew similar conclusions when looking at correlations with other alcohol measures only.

As predicted none of the correlations were particularly high, however it is of note that there is some disparity between the factors regarding physical and mental health (as measured by the SF-12), with mental health being associated with dependence and consequences and physical health being significantly related to consumption patterns. If the scale was treated as unidimensional, there was no significant relationship between physical health and hazardous drinking. These relationships may be worth the attention of further research with more in depth measurement instruments to explore this finding. In addition, the PSQ score was only associated with dependence/consequences but not consumption, further supporting the need to view the AUDIT score as a two dimensional scale. It could be suggested that alternative scoring may improve the ability to distinguish binge drinking and signs of dependence and problems.

Chung *et al*. (2002) suggested that the value of having a multidimensional or unidimensional scale depends on the goals of the instrument. Treating the scale as two dimensions would have benefits regardless of the use of the scale. If the AUDIT is being used in research as a variable

to indicate hazardous drinking, the conceptual differences between consumption and the consequences/dependences of drinking have demonstrated unique predictive validity and relation to different additional and related variables. If the AUDIT is being used as a screening instrument, for example in a primary care setting, creating scores for two factors may marginally increase the time involved in scoring and interpretation. However, it could additionally be argued that excessive consumption, whether or not an individual is experiencing any further alcohol related problems also needs investigation given the disparity between scores on the mental and physical SF-12 components. With this in mind, it may be advantageous to explore factor specific cut points to assist clinicians with administration and interpretation.

It is worthy of comment that the factor loadings for item 9 was very low (0.29). The poor functioning of this item has been reported previously by Karno *et al*. (2000) and Kelly and Donovan (2001). Lima *et al*. (2005) excluded this item because it was not generating score variability.

In conclusion, with mounting evidence that scores derived from the AUDIT are best explained in terms of two correlated dimensions it would appear timely to explore the possibilities of scoring and interpreting the scale in a manner that is consistent with such findings.

REFERENCES

Allen, J. P., Litten, R. Z., Fertig, J. B. and Babor, T. (1997) A review of research on the Alcohol Use Disorders Identification Test (AUDIT). *Alcoholism-Clinical and Experimental Research* **21**, 613–619.

Babor, T. F., de la Fuente, J. R., Saunders, J. and Grant, M. (1992) *AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for Use on Primary Health Care*. World Health Organisation, Geneva.

Bebbington, P. E. and Nayani, T. (1995) The psychosis screening questionnaire. *International Journal of Methods in Psychiatric Research* **5**, 11–19.

Bentler, P. M. (1990) Comparative fit indices in structural models. *Psychological Bulletin* **107**, 238–246.

Bergman, H. and Kallmen, H. (2002) Alcohol use among Swedes and a psychometric evaluation of the alcohol use disorders identification test. *Alcohol and Alcoholism* 37, 245–251.

Bollen, K. A. (1989) *Structural Equations with Latent Variables*. Wiley, New York.

Browne, M. W. and Cudeck, R. (1989) Single sample cross-validation indices for covariation structures. *Multivariate Behavioral Research* 24, 445–455.

Carey, K. B., Carey, M. P. and Chandra, P. S. (2003) Psychometric evaluation of the alcohol use disorders identification test and short drug abuse screening test with psychiatric patients in India. *Journal of Clinical Psychiatry* 64, 767–774.

Chung, T., Colby, S. M., Barnett, N. P. and Monti, P. M. (2002) Alcohol use disorders identification test: factor structure in an adolescent emergency department sample. *Alcoholism-Clinical and Experimental Research* 26, 223–231.

Curran, P. J., West, S. G. and Finch, J. F. (1996) The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods* 1, 16–29.

El-Bassel, N., Schilling, R., Ivanoff, A. *et al*. (1998) Assessing the World Health Organisation's alcohol use disorder identification test among incarcerated women. *Journal of Offender Rehabilitation* 26, 71–89.

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

Hoyle, R. H. and Panter, A. T. (1995) Writing about structural equation models. In *Structural Equation Modeling: Concepts, Issues and Applications*, Hoyle., R. H. ed., pp. 158–198. Sage, London.

Hu, L. and Bentler, P. M. (1998) Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychological Methods* 4, 424–453.

Hu, L. and Bentler, P. M. (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modelling* 6, 1–55.

Jané-Llopis, E. and Matytsina, I. (2006) Mental health and alcohol, drugs and tobacco: a review of the comorbidity between mental disorders and the use of alcohol, tobacco and illicit drugs. *Drug and Alcohol Review* 25, 515–536.

Jenkinson, C. and Layte, R. (1997) Development and testing of the UK SF-12 (short form health survey). *Journal of Health Services Research and Policy* 2, 14–16.

Jöreskog, K. and Sörbom, D. (1981) *LISREL V: Analysis of Linear Structural Relationships by the Method of Maximum Likelihood*. National Educational Resources, Chicago.

Jöreskog, K. G. and Sörbom, D. (1993) *Structural Equation Modeling with the SIMPLIS Command Language*. Scientific Software Inc, Chicago.

Jöreskog, K. and Sörbom, D. (2004a) *LISREL 8.70*. Scientific Software Inc, Chicago.

Jöreskog, K. and Sörbom, D. (2004b) *PRELIS 2.70*. Scientific Software Inc, Chicago.

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

Kelly, T. M. and Donovan, J. E. (2001) Confirmatory factor analyses of the Alcohol Use Disorders Identification Test (AUDIT) among Adolescents Treated in Emergency Departments. *Journal of Studies on Alcohol* 62, 838–842.

Lewis, G. and Pelosi, A. (1990) *Manual of the Revised Clinical Interview Schedule (CIS-R)*. Institute of Psychiatry, London.

Lewis, G., Pelosi, A. J., Araya, R. and Dunn, G. (1992) Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychological Medicine* 22, 465–486.

Lima, C. T., Freire, A. C. C., Silva, A. P. B. *et al*. (2005) Concurrent and construct validity of the AUDIT in an urban Brazilian sample. *Alcohol and Alcoholism* 40, 584–589.

MacDonald, Z. and Shields, M. A. (2001) The impact of alcohol consumption on occupational attainment in England. *Economica* 271, 427–453.

Maisto, S. A., Conigliaro, J., McNeil, M. *et al*. (2000) An empirical investigation of the Factor Structure of the AUDIT. *Psychological Assessment* 12, 346–353.

Medina-Mora, E., Carreño, S. and De la Fuente, J. R. (1998) Experiences with the Alcohol Use Disorders Identification Test (AUDIT) in Mexico. In *Recent Developments in Alcoholism, Vol. 14, The Consequences of Alcoholism*, Galanter., M. ed., pp. 383–396. Plenum Press, New York.

Mortensen, E. L., Jensen, H. H., Sanders, S. A. and Reinisch, J. M. (2006) Associations between volume of alcohol consumption and social status, intelligence and personality in a sample of young adult Danes. *Scandinavian Journal of Psychology* 47, 387–398.

Nelson, H. E. and O'Connell, A. (1978) Dementia. The estimation of pre-morbid intelligence levels using the New Adult Reading Test. *Cortex* 14, 234–244.

Nelson, H. E. and Willison, J. (1991) *National Adult Reading Test (NART) Test Manual*. NFER- Nelson, Windsor, England.

O'Hare, T. and Sherrer, M. V. (1999) Validating the alcohol use disorder identification test with college first-offenders. *Journal of Substance Abuse Treatment* 17, 113–119.

Reinert, D. F. and Allen, J. P. (2002) The Alcohol Use Disorders Identification Test (AUDIT): A review of recent research. *Alcoholism-Clinical and Experimental Research* 26, 272–279.

Satorra, A. (1992) Asymptotic robust inferences in the analysis of mean and covariance structures. *Sociological Methodology* 22, 249–278.

Saunders, J. B. and Aasland, O. G. (1987) *WHO Collaborative Project on Identification and Treatment of Persons with Harmful Alcohol Consumption*, WHO/MNH/DAT 86.3. World Health Organization: Geneva.

Shields, A. L., Guttmanova, K. and Caruso, J. C. (2004) An examination of the factor strucute of the alcohol use disorders identification test in two high-risk samples. *Substance Use & Misuse* 39, 1161–1182.

Singleton, N., Bumpstead, R., O'Brien, M. *et al*. (2001) *Psychiatric Morbidity Among Adults Living in Private Households, 2000*. Stationery Office, London.

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

Steiger, J. H. (1990) Structural model evaluation and modification: an interval estimation approach. *Multivariate Behavioral Research* 25, 173–180.

Tanaka, J. S. (1987) "How big is big enough?" Sample size and goodness of fit in structural equation models with latent variables. *Child development* 58, 134–146.

van Ours, J. C. (2004) A pint a day raises a man's pay; but smoking blows that gain away. *Journal of Health Economics* 23, 863–886.

Ware, J. E., Kosinski, M. and Keller, S. D. (1996) A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Medical Care* 34, 220–233.

Windle, M. and Blane, H. T. (1989) Cognitive ability and drinking behavior in a national sample of young adults. *Alcoholism-Clinical and Experimental Research* 13, 43–48.