The dimensionality of alcohol use disorders: Results from Israel

Dvora Shmulewitz [a,b], Katherine Keyes [b,c], Cheryl Beseler [b,d,e], Efrat Aharonovich [a,b], Christina Aivadyan [a], Baruch Spivak [f], Deborah Hasin [a,b,c,*]

[a] New York State Psychiatric Institute, New York, NY 10032, USA
[b] Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA
[c] Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY 10032, USA
[d] Department of Epidemiology, University of Nebraska Medical Center, Omaha, NE 68198, USA
[e] Department of Environmental, Agricultural and Occupational Health Sciences, Omaha, NE 68198, USA
[f] Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

ARTICLE INFO

ABSTRACT

Aims: To prepare for DSM-V, the structure of DSM-IV alcohol dependence and abuse criteria and a proposed additional criterion, at-risk drinking, require study in countries with low per-capita consumption, and comparison of current and lifetime results within the same sample. We investigated DSM-IV Alcohol Use Disorder (AUD) criteria in Israel, where per-capita alcohol consumption is low.
Methods: Household residents selected from the Israeli population register (N = 1338) were interviewed with the AUDADIS. Item response theory analyses were conducted using MPlus, and diagnostic thresholds were examined with the kappa statistic.
Results: Dependence and abuse criteria fit a unidimensional model interspersed across the severity continuum, for both current and lifetime timeframes. Legal problems were rare and did not improve model fit. Weekly at-risk drinking reflected greater severity than in U.S. samples. When dependence and abuse criteria were combined, a diagnostic threshold of ≥3 criteria produced the best agreement with DSM-IV diagnoses (kappa > 0.80).
Conclusion: Consistent with other studies, alcohol dependence and abuse criteria reflected a latent variable representing a single AUD. Results suggested little effect in removing legal problems and little gained by adding weekly at-risk drinking. Results contribute to knowledge about AUD criteria by examining them in a low-consumption country.

© 2010 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

The Diagnostic and Statistical Manual of Mental Disorder, fourth edition (DSM-IV; American Psychiatric Association, 1994) defines two alcohol use disorders (AUDs), alcohol dependence and abuse. Dependence requires at least three of seven criteria within a 12-month period, while abuse requires at least one of four criteria, in the absence of lifetime dependence. Two assumptions underlie these categories: (1) alcohol dependence and abuse are distinct disorders and (2) alcohol abuse is hierarchically less important than dependence, implying lower severity.

In preparation for DSM-V, these assumptions are under examination, using item response theory (IRT) and other methods. Studies

in the U.S. (Gelhorn et al., 2008; Harford et al., 2009; Hasin and Beseler, 2009; Langenbucher et al., 2004; Martin et al., 2006; Saha et al., 2006, 2007) and elsewhere (Borges et al., 2010; Proudfoot et al., 2006) show two highly correlated factors or a single latent variable, suggesting that dependence and abuse constitute a single condition, with dependence and abuse criteria interspersed across an underlying severity spectrum. This literature contributes important information, but leaves several issues open.

The first issue concerns cross-national generalizability of the alcohol dependence and abuse criteria. Thus far, work was conducted in the United States, Australia, Argentina, Mexico and Poland, all countries with moderately high per-capita total alcohol consumption (7.65–10.94 l/year; World Health Organization, 2009). No information exists on the structure of alcohol dependence and abuse criteria in countries with lower consumption.

The second issue is whether timeframe (current or lifetime) affects the results. Thus far, IRT results on *current* criteria come from general population (Harford et al., 2009; Proudfoot et al., 2006; Saha et al., 2006, 2007) and emergency room (ER) samples (Borges et al.,

2010), while results on *lifetime* criteria come from clinical samples (Gelhorn et al., 2008; Langenbucher et al., 2004; Martin et al., 2006). These studies suggest some differences between the discrimination and severity of lifetime and current criteria. Current criteria may be less subject to recall bias than lifetime criteria (Saha et al., 2006), but lifetime measures are needed for many purposes, including genetic studies. Whether previous discrepancies arose from variation in timeframe or sample type is unknown.

The third issue concerns proposed changes to DSM-V AUD criteria. One proposal is to remove legal problems (Martin et al., 2008), which fit poorly in factor analysis (FA), had low discrimination, and show differential item functioning (Borges et al., 2010; Gelhorn et al., 2008; Harford et al., 2009; Kahler and Strong, 2006; Langenbucher et al., 2004; Martin et al., 2006; Proudfoot et al., 2006; Saha et al., 2006, 2007). The legal problems criterion is rare, and is often endorsed with other items, suggesting minimal effect on AUD prevalence if removed.

Another proposed change involves adding a consumption criterion (Borges et al., 2010; Keyes et al., 2009; Li et al., 2007a,b; Martin et al., 2008; Saha et al., 2007). IRT analysis of current AUD criteria in the U.S. general population suggested that at-risk drinking (5+ drinks/occasion for men; 4+ drinks for women) at least weekly fits with the DSM-IV diagnostic criteria, capturing the milder end of the spectrum (Saha et al., 2007). IRT analyses of current AUD criteria in ER samples suggested that various consumption criteria tap the low-to-middle severity range and perform somewhat differently across countries, but did not directly test weekly at-risk drinking (Borges et al., 2010). Discontinuity analysis of lifetime AUD criteria in the U.S. general population did not support adding weekly at-risk drinking to DSM-V (Hasin and Beseler, 2009). However, IRT analyses that include lifetime AUD criteria and weekly at-risk drinking have not been conducted.

The fourth issue is differential item functioning (DIF). Studies have noted DIF by age and gender for several AUD criteria (Harford et al., 2009; Kahler and Strong, 2006; Martin et al., 2006; Saha et al., 2006). Whether this occurs in other countries is unknown.

The fifth issue involves diagnostic thresholds. If dependence and abuse criteria are combined to diagnose one AUD, a diagnostic threshold will need to be determined. In the absence of a clear discontinuity at some point in the AUD severity spectrum, a threshold producing greatest agreement with DSM-IV prevalence may be the most practical solution. We know of no studies comparing concordance between potential thresholds of AUD criteria and DSM-IV AUD diagnoses in non-U.S. samples.

We addressed these issues in an Israeli sample collected to study genetic and environmental influences on drinking (Hasin et al., 2002a,b; Spivak et al., 2007). Per-capita total alcohol consumption in Israel (3.28 l/year; World Health Organization, 2009) is lower than in other countries in which IRT AUD analyses were conducted; for example, total per-capita consumption is 9.58 l/year (World Health Organization, 2009) in the United States. Among Israeli samples, lifetime abstention is rare but heavy drinking and alcohol problems are relatively uncommon (Baras et al., 1984; Levav et al., 1997; Yeung and Greenwald, 1992). Thus, Israel provides a contrasting context for AUD studies. In addition, the AUD measure (Grant et al., 1995, 2003) used in the Israel study was also used in U.S. IRT studies (Saha et al., 2007), thus holding instrumentation constant. We addressed these questions: (1) do the DSM-IV alcohol dependence and abuse criteria show unidimensionality and dispersion across the severity continuum? (2) Do current and lifetime criteria show similar results within the same sample? (3) Does evidence support removing legal problems or adding weekly at-risk drinking? (4) Is DIF consistent with other studies? (5) What number of criteria could serve as a threshold to diagnose an AUD?

2. Methods

2.1. Study procedures

Adult household residents were selected from the Israeli population register with these characteristics: Jewish ethnicity, native Israeli or immigrant from the Former Soviet Union (FSU), and 3:1 oversampling for males. These characteristics were consistent with the overall research goals of examining the genetic and environmental influences on drinking. Immigrants from the FSU form a comparison group to native Israelis because (a) over 1,000,000 immigrated since 1989, comprising a large group relative to the overall population (7.1 million), and (b) they are more likely to drink, drink frequently, and experience alcohol dependence symptoms than other Israelis (Hasin et al., 1998, 2002a,b; Rahav et al., 1999). We oversampled for males because addressing the study goals would have been difficult in Israeli women, who have very low rates of drinking (Baras et al., 1984; Neumark et al., 2001). Data collection followed previous procedures (Hasin et al., 2002a,b; Spivak et al., 2007). Potential respondents received an introductory letter and follow-up call to explain the study and schedule an in-person interview. Interviewers obtained written informed consent as approved by institutional review boards at New York State Psychiatric Institute, Tel Aviv University, Ness Ziona/Ba'er Yaakov Psychiatric Hospitals and the Israel Ministry of Health. Interviewers administered computer-assisted interviews in Hebrew or Russian. Among eligible participants, the response rate was 68.9%.

Interviewers were health professionals or survey interviewers who underwent 6 days of training via a manual, self-study exercises, didactic presentations, role-plays, and final supervisor certification. Ongoing supervision was conducted by periodic supervisor field observation, structured review of recorded interviews, and telephone verification of participation and responses to questions.

2.2. Sample

The sample for the FA and IRT analysis included lifetime drinkers ($\geq$1 drink ever, $N = 1160$), of whom 80.5% ($N = 934$) were male, 25.6% ($N = 297$) were immigrants from the FSU, and 91.9% ($N = 1066$) were current drinkers ($\geq$1 drink in prior 12 months). Age was dichotomized by the median (38 years), with 49.7% ($N = 576$) in the younger group. The agreement analysis (see Section 2.4.3 below) was conducted on the entire sample, $N = 1338$, since we compared *overall* prevalence between DSM-IV and potential DSM-V diagnoses.

2.3. Alcohol measures and criteria

Alcohol dependence and abuse and criteria were measured with the Alcohol Use Disorders and Associated Disabilities Interview Schedule (AUDADIS; Grant et al., 1995, 2003), used in U.S. (Grant, 1997; Grant et al., 2004; Hasin et al., 1997a, 2007) and international studies (Canino et al., 1999; Chatterji et al., 1997; Vrasti et al., 1998). AUDADIS methods to assess current and lifetime DSM-IV criteria are covered in detail elsewhere (Grant et al., 1995; 2004; Hasin et al., 2007). Diagnoses of DSM-IV alcohol dependence and abuse were made consistent with the National Epidemiologic Survey on Alcohol and Related Conditions (Grant et al., 2004; Hasin et al., 2007). Specifically, DSM-IV dependence was diagnosed by the presence of three or more of seven dependence criteria within a 12-month timeframe that cause significant distress or impairment, while DSM-IV abuse was diagnosed by the presence of one or more of four recurrent abuse criteria within a 12-month timeframe that cause significant distress or impairment, in the absence of lifetime alcohol dependence. The seven dependence and four abuse criteria are listed in Tables 1 and 2.

Reliability of AUDADIS-IV alcohol diagnoses in clinical and general population samples (Chatterji et al., 1997; Grant et al., 1995, 2003; Hasin et al., 1997a) ranges from very good-to-excellent ($K = 0.70$–0.84). Convergent, discriminative and construct validity of AUDADIS-IV AUD criteria and diagnoses is good-to-excellent in U.S. (Hasin and Paykin, 1999; Hasin et al., 1997c) and international studies (Hasin et al., 1997b; Pull et al., 1997; Vrasti et al., 1998), including concordance of AUDADIS diagnoses with clinician reappraisals ($K = 0.60$–0.76; Canino et al., 1999; Cottler et al., 1997).

Respondents were asked about quantity and frequency of consumption of wine, beer, liquor and pre-mixed drinks in the last 12 months, and if relevant, total drinking during the heaviest prior drinking period. The consumption items have good psychometric properties (Grant et al., 1995, 2003; Hasin et al., 1997b). After converting to standard drink sizes, we created an at-risk drinking criterion (5+ drinks for males, 4+ drinks for females at least weekly) consistent with NIAAA guidelines for current and lifetime timeframes.

2.4. Statistical analysis

2.4.1. Analysis of internal consistency and dimensionality. Internal consistency of the alcohol criteria was measured by Cronbach's coefficient α (Cronbach, 1951), using SAS 9.1.3 (SAS Institute Inc., 2008). Factor analyses and IRT analyses were conducted with MPlus (Muthén and Muthén, 2008). We used exploratory and confirmatory factor analyses to test unidimensionality. Number of factors was determined with the Kaiser criterion (recommended value of eigenvalues to retain factors, >1.0) and Cattell's Scree Test (retaining factors appearing on a graph of eigenvalues before the

Table 1
Factor analysis and item response theory for current drinkers, $N = 1066$.

Alcohol criteria[a]	Prevalence %	Factor loadings	Item response theory parameters	
			Discrimination (s.e.)	Severity (s.e.)
Dependence				
Tolerance	19.9	0.786	1.364 (0.16)	1.066 (0.07)
Withdrawal	6.9	0.729	1.239 (0.16)	1.946 (0.14)
Larger/longer	15.9	0.826	1.600 (0.20)	1.201 (0.07)
Quit/control	8.4	0.661	1.003 (0.12)	1.987 (0.15)
Time spent	4.1	0.751	1.349 (0.19)	2.199 (0.16)
Activities given up	0.9	0.845	2.011 (0.45)	2.624 (0.22)
Physical/psychological	8.3	0.652	0.900 (0.12)	2.114 (0.18)
Abuse				
Neglect roles	1.4	0.846	1.920 (0.49)	2.481 (0.21)
Hazardous use	8.0	0.667	0.999 (0.13)	2.028 (0.16)
Legal problems	0.4	0.641	1.239 (0.35)	3.606 (0.62)
Social/interpersonal	2.7	0.856	1.881 (0.36)	2.190 (0.15)
Weekly at-risk drinking[b]	10.5	0.719	1.081 (0.13)	1.735 (0.12)

Model fit indices[c]

Factor analysis	Item response theory						
	DSM-IV criteria	Without Legal problems	Including weekly at-risk drinking		DSM-IV criteria	Legal Problems, constrained	Weekly at-risk drinking, free	Weekly at-risk drinking, constrained
CFI[d]	0.995	0.992	0.993	AIC[e]	4486.272	4496.449	5059.029	5204.780
TLI[d]	0.995	0.993	0.994	BIC[e]	4595.648	4600.854	5178.349	5319.129
RMSEA[e]	0.015	0.019	0.016	SS-BIC[e]	4525.772	4534.154	5102.121	5246.077

[a] Description of DSM-IV criteria (American Psychiatric Association, 1994).
Tolerance: diminished effect or need more to get desired effect.
Withdrawal: withdrawal syndrome or alcohol or related substance taken to prevent or relieve symptoms.
Larger/longer: drinking more or for longer then wanted to.
Quit/control: wanted to stop or drink less, or unsuccessful attempts.
Time spent: spent a lot of time getting alcohol, drinking, or recovering from drinking.
Activities given up: gave up important activities to drink.
Physical/psychological: continued drinking despite knowledge that it caused or exacerbated a physical or psychological problem.
Neglect roles: drinking caused failure to conduct work, home, or school obligations.
Hazardous use: drinking that incurred danger, e.g., driving.
Legal problems: legal problems due to drinking.
Social/interpersonal: drinking despite knowledge that it caused or exacerbated social or interpersonal problems.
[b] Weekly at-risk drinking: weekly or more frequent consumption of 5 or more drinks per occasion for men, 4 or more for women.
[c] CFI = comparative fit index, TLI = Tucker–Lewis index, RMSEA = root mean square error of approximation, AIC = Akaike information criterion, BIC = Bayesian information criterion, SS-BIC = sample-size corrected BIC.
[d] Larger value indicates better model fit.
[e] Smaller value indicates better model fit.

eigenvalue curve begins to flatten) (Norman and Streiner, 2008). Comparative Fit Index (CFI) ≥ 0.95, Tucker–Lewis Index (TLI) ≥ 0.95, and root mean square error of approximation (RMSEA) ≤ 0.06 (Hu and Bentler, 1999) were used to assess model fit.

To do an IRT analysis in MPlus, first a confirmatory factor analysis is conducted, using robust maximum likelihood estimation (Baker and Kim, 2004; Muthén and Muthén, 1998–2007). The resulting factor estimates (factor loadings and thresholds) are converted to IRT estimates (item discrimination and severity) using a two-parameter logistic IRT model for dichotomous traits (Birnbaum, 1968; Lord and Novick, 1968). This model estimates the probability of endorsing an item at any value of the underlying latent trait as a function of item discrimination and severity. Discrimination, related to factor loading, is the ability to discriminate among respondents with latent trait levels above or below the item severity. Severity, inversely related to item prevalence, is the latent trait value that confers 50% probability of endorsing that item.

Item Characteristic Curves (ICC) show severity, the point on the x-axis where the y-axis is 0.5, and discrimination, the slope of the curve at that point. Item Information Curves (IIC) show the amount of information the item provides across the latent trait continuum; information peaks at item severity, and peak height is directly related to discrimination (Krueger et al., 2004). ICCs and IICs are shown to enable visualization of the relationships between item severity, discrimination/information, and the underlying latent trait. Total Information Curves (TIC) show total test information, calculated by summing the information for all items for all latent trait values.

We calculated Cronbach's α and conducted factor and IRT analysis using current (past 12 months) criteria in current drinkers and lifetime criteria in lifetime drinkers. We ranked items by discrimination and severity order and descriptively compared item ranking across timeframes.

Analyses were done with and without legal problems and weekly at-risk drinking, comparing CFI, TLI, and RMSEA to see if their presence affected unidimensionality. To assess IRT model fit, we compared the Akaike Information Criterion

(AIC), Bayesian Information Criterion (BIC) and sample-size corrected BIC (SS-BIC) between models constraining the criterion (legal problems or weekly at-risk drinking) loadings to 0 or freeing the criterion loadings to take any value; lower values indicate better model fit. There are no standard statistical methods to judge the extent or significance of fit improvement based on the reduction in the AIC, BIC, or SS-BIC. Instead, there are statistical tests available to compare the likelihoods of these nested models, such as the likelihood ratio test (LRT). We used the LRT to compare the nested models in which criteria were fixed to have a loading of 0 versus free to vary. The LRT uses -2 times the difference in the loglikelihoods of nested models to calculate an adjusted χ^2-statistic to determine which model shows better fit, and the size and significance of the improvement in the model fit (Muthén and Muthén, 2009).

2.4.2. Differential item functioning analysis. The LRT was used to identify lifetime criteria with differential item functioning (DIF; Thissen et al., 1988, 1993) by gender, being an immigrant from the FSU, and age. These variables are known to influence drinking patterns in Israel (Hasin et al., 1998, 2002a,b; Rahav et al., 1999). A series of LRTs were done, comparing models with item estimates constrained as equal across groups to models allowing estimates to vary (Strong et al., 2009; Thissen, 2001). For each item, we compared Model 1, a fully constrained model with factor estimates (loadings and thresholds) for all items equal in each group, to Model 2, where factor estimates can vary for that item only. If this LRT is significant, the item shows DIF (allowing estimates to differ by group improves model fit). Then, for each item with DIF, we compared Model 3 (constraining the loadings to be equal across groups but allowing thresholds to vary) to Model 2. A significant LRT demonstrates a difference in the slope or discrimination (allowing loadings to vary improves model fit). Next, we compared Models 1 and 3. Here, a significant LRT demonstrates a difference in severity (allowing thresholds to vary, conditional on equal slopes, improves model fit). Latent trait severity had a mean of 0 and variance of 1 in the reference group (female, non-FSU group, older age); mean and variance were calculated along with

Table 2
Factor analysis and item response theory for lifetime drinkers, $N = 1160$.

Alcohol criteria[a]	Prevalence %	Factor loadings	Item response theory parameters	
			Discrimination (s.e.)	Severity (s.e.)
Dependence				
Tolerance	35.3	0.734	1.128 (0.10)	0.513 (0.06)
Withdrawal	16.1	0.655	0.923 (0.09)	1.472 (0.10)
Larger/longer	36.8	0.776	1.269 (0.12)	0.437 (0.05)
Quit/control	13.7	0.615	0.815 (0.09)	1.748 (0.14)
Time spent	11.6	0.768	1.288 (0.14)	1.535 (0.09)
Activities given up	2.0	0.871	1.826 (0.41)	2.379 (0.18)
Physical/psychological	11.8	0.716	1.099 (0.11)	1.623 (0.10)
Abuse				
Neglect roles	2.9	0.914	2.409 (0.46)	2.052 (0.11)
Hazardous use	17.7	0.687	1.022 (0.10)	1.310 (0.09)
Legal problems	2.3	0.684	1.127 (0.19)	2.753 (0.26)
Social/interpersonal	10.0	0.812	1.492 (0.17)	1.560 (0.09)
Weekly at-risk drinking[b]	15.9	0.701	1.045 (0.10)	1.400 (0.09)

Model fit indices[c]

Factor analysis	Item response theory						
	DSM-IV criteria	Without legal problems	Including weekly at-risk drinking		DSM-IV criteria	Legal Problems, constrained	Weekly at-risk drinking, free	Weekly at-risk drinking, constrained
CFI[d]	0.995	0.995	0.990	AIC[e]	7822.412	7881.864	8620.227	8839.123
TLI[d]	0.997	0.996	0.993	BIC[e]	7933.648	7988.044	8741.576	8955.415
RMSEA[e]	0.016	0.018	0.023	SS-BIC[e]	7863.768	7921.341	8665.344	8882.359

[a] Description of DSM-IV criteria (American Psychiatric Association, 1994).
Tolerance: diminished effect or need more to get desired effect.
Withdrawal: withdrawal syndrome or alcohol or related substance taken to prevent or relieve symptoms.
Larger/longer: drinking more or for longer then wanted to.
Quit/control: wanted to stop or drink less, or unsuccessful attempts.
Time spent: spent a lot of time getting alcohol, drinking, or recovering from drinking.
Activities given up: gave up important activities to drink.
Physical/psychological: continued drinking despite knowledge that it caused or exacerbated a physical or psychological problem.
Neglect roles: drinking caused failure to conduct work, home, or school obligations.
Hazardous use: drinking that incurred danger, e.g., driving.
Legal problems: legal problems due to drinking.
Social/interpersonal: drinking despite knowledge that it caused or exacerbated social or interpersonal problems.
[b] Weekly at-risk drinking: weekly or more frequent consumption of 5 or more drinks per occasion for men, 4 or more for women.
[c] CFI = comparative fit index, TLI = Tucker–Lewis index, RMSEA = root mean square error of approximation, AIC = Akaike information criterion, BIC = Bayesian information criterion, SS-BIC = sample-size corrected BIC.
[d] Larger value indicates better model fit.
[e] Smaller value indicates better model fit.

item estimates for the focal group (male, FSU group, younger age). We employed Bonferroni corrections for multiple testing. For criteria with significant DIF, we calculated the probability of endorsing the criterion within each group, across all values of the latent trait. These probabilities provide an indication of the magnitude of the DIF effect.

2.4.3. Agreement analysis. To investigate potential diagnostic AUD thresholds for the combined dependence and abuse criteria, we computed kappa (chance-corrected agreement) between a given threshold for DSM-V (e.g., ≥2 criteria) and DSM-IV diagnoses of dependence or abuse, for current and lifetime timeframes, using SAS 9.1.3 (SAS Institute Inc., 2008).

3. Results

3.1. Current diagnostic criteria

In current drinkers ($N = 1066$), current AUD criteria ranged in prevalence from 19.9% (tolerance) to 0.4% for legal problems (Table 1). Current AUD items were internally consistent ($\alpha = 0.74$) and fit a unidimensional model, with significant loadings (0.85 [neglect roles]–0.64 [legal problems]) on a single factor. First and second eigenvalues (6.49 and 0.97), and a scree plot that began leveling at the second factor supported unidimensionality, as did fit indices (Table 1; CFI and TLI = 0.995; RMSEA = 0.015).

IRT parameters and ICC are shown in the upper portion of Table 1 and Fig. 1, respectively. The order of criterion severity, from low to high, was: tolerance, larger/longer, withdrawal, quit/control,

hazardous use, physical/psychological, social/interpersonal, time spent, neglect roles, activities given up, and legal problems.

Item information (Fig. 2), from high to low, was as follows: activities given up, neglect roles, social/interpersonal, larger/longer, tolerance, time spent, legal problems, withdrawal, quit/control, hazardous use and physical/psychological. Fig. 2 (inset) shows the TIC, summed for all DSM-IV criteria. Total information was highest for the mid-to-high severity range, with maximum information (12.22) at 2.4.

3.2. Lifetime diagnostic criteria

The prevalence of lifetime AUD criteria in lifetime drinkers ($N = 1160$) ranged from 36.8% for larger/longer to 2.0% for activities given up (Table 2). Lifetime criteria were internally consistent ($\alpha = 0.79$) and loaded significantly (0.91 [neglect roles]–0.62 [quit/control]) on a single factor. A one-factor model was determined by first and second eigenvalues of 6.55 and 0.80, supported by model fit indices (Table 2; CFI = 0.995; TLI = 0.997; RMSEA = 0.016).

IRT parameters are shown in Table 2 (upper portion) and ICCs in Fig. 3. Compared to current criteria, all lifetime criteria had lower severity, reflecting higher prevalence. The order of item severity was similar, except that quit/control was more severe (by rank order) for the lifetime than current timeframe.
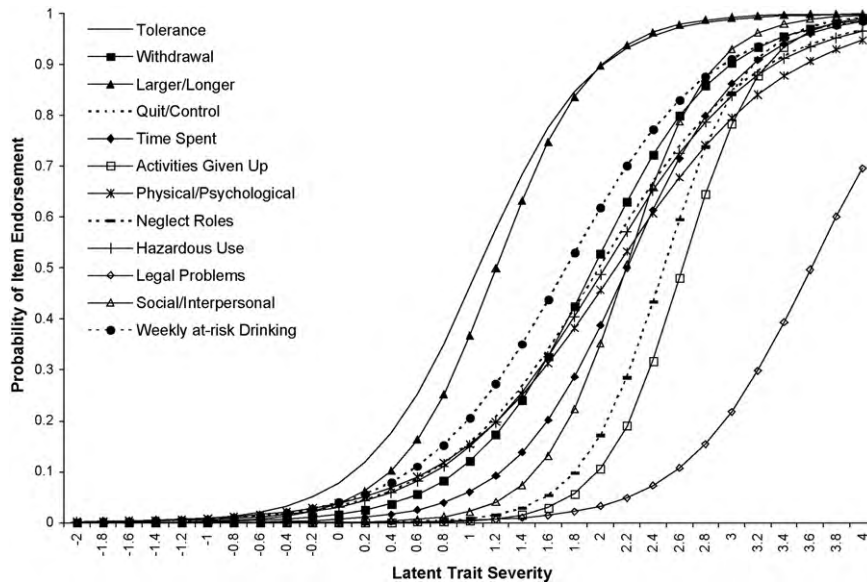
Fig. 1. Item characteristic curves, current drinkers.

Item information for lifetime criteria (Fig. 4) differed slightly from current criteria, with the following order: neglect roles, activities given up, social/interpersonal, time spent, larger/longer, legal problems, tolerance, physical/psychological, hazardous use, withdrawal, and quit/control. Similar to current criteria, total information was sizeable for the mid-to-high severity range, with maximum information (11.22) at 2 (Fig. 4, inset).

3.3. Legal problems

FA fit indices for models with and without legal problems were essentially the same (current, Table 1; lifetime, Table 2). IRT fit indices for models with legal problems constrained were larger

(worse fit) than models with legal problems unconstrained (lower portions of Tables 1 and 2) and the likelihood ratio test (LRT) indicated better fit for the unconstrained models (current, $\chi^2_1 = 17.93$, $p < 0.0001$; lifetime, $\chi^2_1 = 55.71$, $p < 0.0001$). This criterion added little information (see TICs, insets to Figures 2 and 4).

3.4. Weekly at-risk drinking

In contrast to U.S. samples, weekly at-risk drinking was not the most prevalent item in either timeframe (Tables 1 and 2). In factor analyses, loadings for AUD criteria were unaffected by the addition of weekly at-risk drinking. With weekly at-risk drinking, model fit
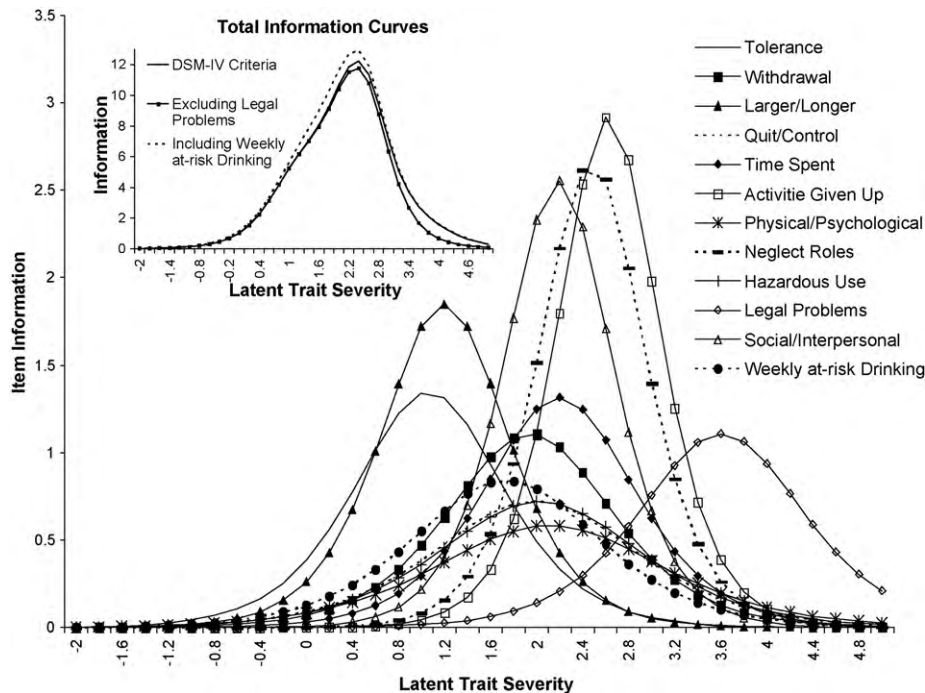

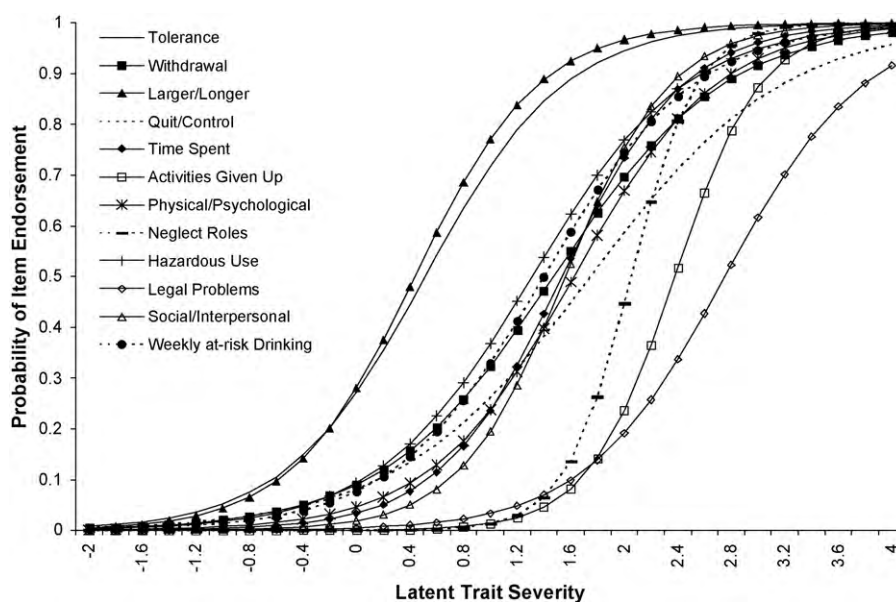**Fig. 2.** Item information curves, current drinkers.

Fig. 3. Item characteristic curves, lifetime Drinkers.

indices were very similar for current and slightly worse for lifetime (lower portions of Tables 1 and 2).

IRT indices for models with weekly at-risk drinking constrained were larger (worse fit) than models with weekly at-risk drinking

unconstrained (lower portions of Tables 1 and 2) and LRTs indicated better fit for the unconstrained models (current, $\chi_1^2 = 151.69$, $p < 0.0001$; lifetime, $\chi_1^2 = 198.29$, $p < 0.0001$). IRT estimates for AUD criteria were similar with and without weekly at-risk drinking.

Table 3
Agreement between DSM-IV alcohol use disorder diagnoses and possible DSM-V diagnoses across a range of potential diagnostic thresholds, N = 1338.

	Current		Lifetime	
	Prevalence %	Kappa[a] (s.e.)	Prevalence %	Kappa[a] (s.e.)
DSM-IV dependence or abuse diagnosis	11.43	–	23.47	–
Potential DSM-V thresholds[b]				
Dependence and abuse criteria				
One or more criteria	27.80	0.50 (0.03)	50.97	0.46 (0.02)
Two or more criteria	15.02	0.71 (0.03)	33.71	0.66 (0.02)
Three or more criteria[c]	8.59	0.84 (0.02)	21.60	0.81 (0.02)
Four or more criteria	4.63	0.55 (0.04)	13.23	0.64 (0.03)
Five or more criteria	2.62	0.34 (0.04)	8.15	0.45 (0.03)
Six or more criteria	1.05	0.15 (0.04)	4.78	0.28 (0.03)
Seven or more criteria	0.67	0.10 (0.03)	2.91	0.18 (0.03)
Eight or more criteria	0.45	0.07 (0.03)	1.64	0.10 (0.02)
Nine or more criteria	0.22	0.03 (0.02)	1.20	0.08 (0.02)
Ten or more criteria	0.07	0.01 (0.01)	0.60	0.04 (0.01)
Dependence and abuse criteria except legal problems				
One or more criteria	27.80	0.50 (0.03)	50.97	0.46 (0.02)
Two or more criteria	15.02	0.71 (0.03)	33.33	0.65 (0.02)
Three or more criteria[c]	8.52	0.84 (0.03)	21.52	0.81 (0.02)
Four or more criteria	4.56	0.54 (0.04)	12.86	0.62 (0.03)
Five or more criteria	2.54	0.34 (0.04)	8.07	0.45 (0.03)
Six or more criteria	1.05	0.15 (0.04)	4.63	0.27 (0.03)
Seven or more criteria	0.67	0.10 (0.03)	2.69	0.17 (0.02)
Eight or more criteria	0.45	0.07 (0.03)	1.57	0.10 (0.02)
Nine or more criteria	0.15	0.03 (0.02)	0.97	0.06 (0.02)
Ten criteria	0.07	0.01 (0.01)	0.22	0.01 (0.01)
Dependence and abuse criteria including weekly at-risk drinking				
One or more criteria	29.67	0.44 (0.03)	51.94	0.44 (0.02)
Two or more criteria	16.29	0.66 (0.03)	35.65	0.65 (0.02)
Three or more criteria[c]	10.01	0.82 (0.03)	23.32	0.79 (0.02)
Four or more criteria	6.13	0.68 (0.04)	15.55	0.72 (0.02)
Five or more criteria	3.29	0.44 (0.04)	10.31	0.55 (0.03)
Six or more criteria	1.87	0.27 (0.04)	6.50	0.37 (0.03)
Seven or more criteria	0.82	0.13 (0.04)	3.89	0.23 (0.03)
Eight or more criteria	0.67	0.11 (0.03)	2.39	0.15 (0.02)
Nine or more criteria	0.45	0.07 (0.03)	1.49	0.09 (0.02)
Ten or more criteria	0.22	0.04 (0.02)	0.52	0.03 (0.01)

[a] Indicates agreement between DSM-IV abuse or dependence diagnosis and a "diagnosis" made at the indicated threshold, e.g. one or more, etc.
[b] Indicates the number of criteria required to make a potential DSM-V diagnosis.
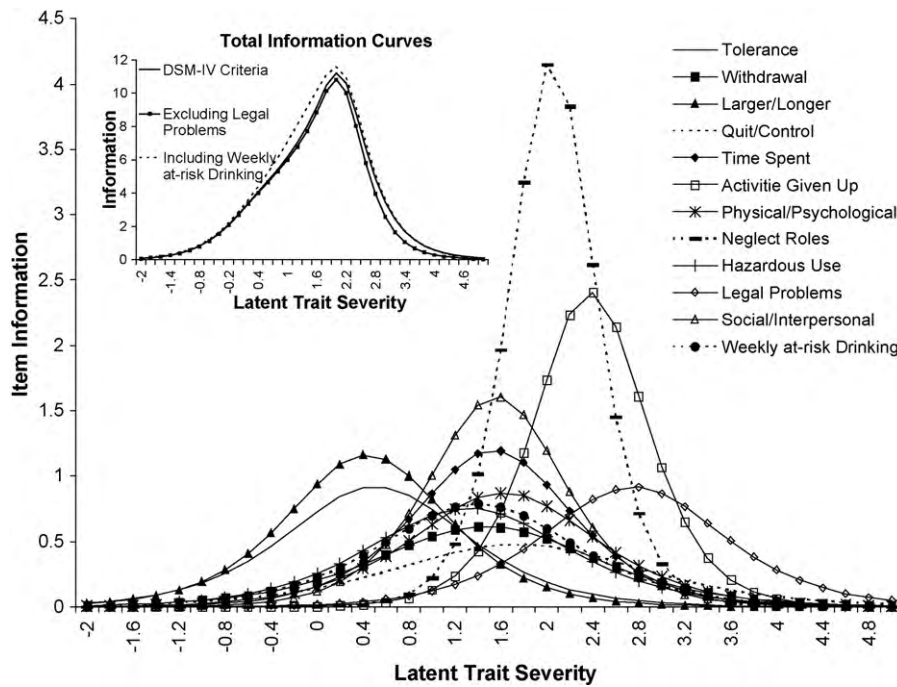[c] Shaded to indicate the highest kappa values.

Fig. 4. Item information curves, lifetime drinkers.

Weekly at-risk drinking showed low to intermediate severity (third from lowest in current, fourth from lowest in lifetime) and low discrimination (fourth from lowest for current and lifetime), but increased total information, particularly for current items (see TICs, insets to Figs. 2 and 4).

3.5. Lifetime differential item functioning

Significant DIF (in the severity parameter) was found by gender and FSU group status, but not by age. For all criteria with DIF, large differences were seen in the probabilities of endorsing the criterion by group among those with the same value of the latent trait; three examples are given here. Social/interpersonal ($\chi^2_1 = 12.54$, $p = 0.0004$) was less severe in males (more likely to be endorsed by males at the same latent trait level). Specifically, at an AUD latent trait level of 1.0, 5.9% of males endorsed social/interpersonal, compared to 0.9% of females, and at a latent trait level of 2.0, 52.7% of males endorsed this criterion, compared to 14.3% of females. Compared to the non-FSU group, tolerance ($\chi^2_1 = 23.52$, $p < 0.0001$) and time spent ($\chi^2_1 = 21.49$, $p < 0.0001$) were more severe in the FSU immigrant group (less likely to be endorsed at the same latent trait level). Specifically, at a latent trait level of 0.0, 11.9% of the FSU group endorsed tolerance, compared to 27.3% of the non-FSU group, and at a trait level of 1.0, 51.5% of the FSU group endorsed this criterion, compared to 74.7% of the non-FSU group. At a latent trait level of 1.0, 8.3% of the FSU group endorsed time spent, compared to 25.8% of the non-FSU group, and at a latent trait level of 2.0, 48.2% of the FSU group endorsed this criterion, compared to 78.0% of the non-FSU group. Physical/psychological ($\chi^2_1 = 11.75$, $p = 0.0006$), legal problems ($\chi^2_1 = 11.57$, $p = 0.0007$), social/interpersonal ($\chi^2_1 = 18.55$, $p < 0.0001$), and weekly at-risk drinking ($\chi^2_1 = 12.40$, $p = 0.0004$) were less severe in the FSU group (more likely to be endorsed at the same latent trait levels) than in the others, with similar levels of divergence between the groups in the likelihood of endorsing the specific criterion at a given value on the AUD latent trait to those given above (available on request).

3.6. Agreement analysis

Among the entire sample ($N = 1338$), the current and lifetime prevalence of DSM-IV alcohol dependence or abuse was 11.43% and 23.47%, respectively (Table 3). For both timeframes, a diagnostic threshold of ≥3 of the 11 combined dependence and abuse criteria agreed best with DSM-IV diagnoses (kappa 0.84, current; 0.81, lifetime). Kappas for immediately adjacent thresholds (≥2 or ≥4) were considerably lower (0.55–0.71). Results were nearly identical when legal problems were removed. When weekly at-risk drinking was added, the same threshold (≥3) agreed best (Table 3; kappa 0.82, current; 0.79, lifetime).

4. Discussion

Consistent with prior studies, factor and IRT analyses of the DSM-IV AUD criteria showed a unidimensional model and dependence and abuse criteria were intermixed across the latent severity continuum. This was true for both the current and the lifetime timeframes. The rank order of dependence and abuse criterion severity was similar for both timeframes. Removal of legal problems slightly improved model fit, reduced AUD prevalence only slightly, and had little other effect. Addition of weekly at-risk drinking slightly worsened model fit although it added some information. Differential item functioning was found mainly among the immigrants from the FSU relative to the rest of the sample, although in both directions.

The unidimensionality and criterion dispersion across the severity spectrum indicate robustness of the AUD structure previously identified, even in countries where per-capita alcohol consumption is low. The main differences between Israeli and U.S. current drinkers (Saha et al., 2007) were found in tolerance and weekly at-risk drinking; tolerance was more severe in the U.S., while weekly at-risk drinking was more severe in Israel, suggesting that tolerance and weekly at-risk drinking are more influenced by national context than other criteria. Furthermore, weekly at-risk drinking exhibited DIF by FSU immigrant status; this item was also more

severe among native Israelis compared to immigrants from the FSU, indicating that nation-of-origin may influence endorsement of this item.

In the present sample, current and lifetime criteria performed similarly in IRT analysis. Although criteria and total information showed higher severity in current drinkers due to lower criterion prevalence, severity order and total test information were similar in both timeframes. Only one exception was found: quit/control was more severe (by rank) in the lifetime than current timeframe, consistent with other studies of current (Harford et al., 2009; Proudfoot et al., 2006; Saha et al., 2006, 2007) and lifetime criteria (Gelhorn et al., 2008; Langenbucher et al., 2004; Martin et al., 2006).

Our findings supported the suggestion to remove legal problems from the diagnostic criteria (Martin et al., 2008). This item was rare, did not improve model fit, added little information, and excluding it had no effect on the agreement analysis. These findings suggest that legal problems contribute little to the total set of diagnostic criteria.

Recent papers suggested adding a consumption criterion (weekly at-risk drinking) to DSM-V (Keyes et al., 2009; Li et al., 2007a,b; Martin et al., 2008; Saha et al., 2007). However, in the Israel sample, weekly at-risk drinking did not improve single-factor model fit, added only slightly to the total test information, and showed low discrimination. Further, weekly at-risk drinking was more severe in Israel than in the U.S. general population sample, thus not meeting the purpose of tapping the mildest end of the continuum (Saha et al., 2007). Similar to the ER study showing that consumption criteria showed DIF across countries (Borges et al., 2010), weekly at-risk drinking showed DIF by FSU status. Alternative definitions of risk drinking (lower frequencies or quantities) did not improve FA or IRT model fit (not shown). These results suggest caution in adding weekly at-risk drinking to the DSM-V AUD criteria.

Several lifetime criteria showed DIF between covariate groups, after accounting for overall differences in latent trait severity. These findings on DIF were not due to trivial differences that were significant only due to the sample-size, but rather reflected substantial group differences in the likelihood of criterion endorsement at the same level of the underlying AUD latent trait. Similar to other studies, males were more likely to endorse social/interpersonal problems (Harford et al., 2009; Kahler and Strong, 2006; Martin et al., 2006; Saha et al., 2006), possibly reflecting antisociality, which is more prevalent among males. FSU immigrants were less likely to endorse tolerance and time spent, and more likely to endorse physical/psychological, legal problems, and social/interpersonal problems at the same level of the latent trait, consistent with findings that their alcohol consumption patterns differ from other Israeli ethnic groups (Hasin et al., 1998, 2002a,b; Rahav et al., 1999). However, the opposite directions of the DIF may balance out, as has been suggested for the DIF found in U.S. samples (Saha et al., 2006, 2007).

DSM-V will need a diagnostic threshold for the continuous AUD trait that minimizes practical difficulties. A threshold maximizing agreement with DSM-IV diagnoses may be best. In the Israeli dataset, the threshold with best agreement was ≥3 AUD criteria. These and results from other datasets, when available, can assist in determining a diagnostic threshold for substance use disorders.

Our results suggest informative phenotypes for genetic studies. For genetic analysis, quantitative traits have more power than that of binary traits, important with limited sample sizes, rare alleles, and/or assessment of gene-gene or gene-environment interaction. Quantitative traits related to AUDs should correlate with the diagnoses and indicate severity (Almasy, 2003). We previously studied alcohol dependence criterion as such a quantitative trait (Hasin et al., 2002a). Results that dependence and abuse criteria comprise one factor justify a quantitative phenotype based on summing all 11 criteria. Further, the number of criteria endorsed and their severity are so highly correlated that additional weighting information based on criterion severity is redundant (Dawson et al., 2010).

We note study limitations. Data were collected by self-report rather than observation, similar to other epidemiologic studies. However, we used a measure with good-to-excellent test–retest reliability that has been well validated. While lifetime criteria may be affected by memory and recall bias, the similarity of lifetime to current results, where recall issues are less salient, suggests that recall problems did not have a strong influence on the lifetime findings.

Study strengths are also noted. The measure, AUDADIS, has excellent reliability and validity, and was used previously in large studies. Data collection involved stringent quality assurance procedures. Consumption of specific drink types over a specific timeframe (current year or past heaviest period) was queried to improve recall (Greenfield and Kerr, 2008). To reduce classification error, at-risk drinking variables were created from frequency and quantity items instead of single items on drinking over the threshold (Dawson, 2003; Greenfield and Kerr, 2008). We provided specific information on the implications of the DIF findings, which other papers have lacked. Finally, the sample contributes unique information on AUD criteria from a low-consumption country to the existing literature.

In summary, IRT results on alcohol dependence and abuse criteria from this Israeli sample show many consistencies and some differences with other IRT studies. In terms of dependence and abuse, both current and lifetime criteria formed a unidimensional construct, interspersed across a severity continuum, suggesting validity even in countries with very different drinking patterns. Based on present results, legal problems could be removed from DSM-V with virtually no effects on overall criteria performance. However, weekly at-risk drinking performed differently in the Israeli sample, suggesting caution in adding this criterion to DSM-V alcohol use disorders. The findings provide important information for DSM-V revisions, and for investigators wishing to use a unidimensional AUD severity construct in epidemiologic or genetic studies.

Role of funding source

Contributors

Authors Aharonovich, Aivadyan, Hasin, and Spivak were involved in data collection and management. Authors Hasin and Shmulewitz designed the study. Authors Beseler, Keyes and Shmulewitz undertook the statistical analysis, and, along with author Hasin, wrote the manuscript. All authors contributed to and have approved the final manuscript.

Conflict of interest

None of the researchers has any connection with the tobacco, alcohol, pharmaceutical or gaming industries or any body substantially funded by one of these organizations.

References

Almasy, L., 2003. Quantitative risk factors as indices of alcoholism susceptibility. Ann. Med. 35, 337–343.

American Psychiatric Association, 1994. Diagnostic and Statistical Manual of Mental Disorder, fourth ed. American Psychiatric Association, Washington, D.C.

Baker, F.B., Kim, S., 2004. Item Response Theory: Parameter Estimation Techniques, second ed. Marcel Dekker, Inc, New York.

Baras, M., Harlap, S., Eisenberg, S., 1984. Alcohol drinking in Jerusalem. Alcohol 1, 435–439.

Birnbaum, A., 1968. Some latent trait models. In: Lord, F.M., Novick, M.R. (Eds.), Statistical Theory of Mental Test Scores. Addison-Wesley, Reading, M.A., pp. 397–472.

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

Canino, G., Bravo, M., Ramirez, R., Febo, V.E., Rubio-Stipec, M., Fernandez, R.L., Hasin, D., 1999. The Spanish Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability and concordance with clinical diagnoses in a Hispanic population. J. Stud. Alcohol 60, 790–799.

Chatterji, S., Saunders, J.B., Vrasti, R., Grant, B.F., Hasin, D., Mager, D., 1997. Reliability of the alcohol and drug modules of the Alcohol Use Disorder and Associated disabilities interview schedule—alcohol/drug-revised (AUDADIS-ADR): an international comparison. Drug Alcohol Depend. 47, 171–185.

Cottler, L.B., Grant, B.F., Blaine, J., Mavreas, V., Pull, C., Hasin, D., Compton, W.M., Rubio-Stipec, M., Mager, D., 1997. Concordance of DSM-IV alcohol and drug use disorder criteria and diagnoses as measured by AUDADIS-ADR, CIDI and SCAN. Drug Alcohol Depend. 47, 195–205.

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

Dawson, D.A., 2003. Methodological issues in measuring alcohol use. Alcohol Res. Health 27, 18–29.

Dawson, D.A., Saha, T.D., Grant, B.F., 2010. A multidimensional assessment of the validity and utility of alcohol use disorder severity as determined by item response theory models. Drug Alcohol Depend. 107, 31–38.

Gelhorn, H., Hartman, C., Sakai, J., Stallings, M., Young, S., Rhee, S.H., Corley, R., Hewitt, J., Hopfer, C., Crowley, T., 2008. Toward DSM-V: an item response theory analysis of the diagnostic process for DSM-IV alcohol abuse and dependence in adolescents. J. Am. Acad. Child Adolesc. Psychiatry 47, 1329–1339.

Grant, B.F., 1997. Prevalence and correlates of alcohol use and DSM-IV alcohol dependence in the United States: results of the National Longitudinal Alcohol Epidemiologic Survey. J. Stud. Alcohol 58, 464–473.

Grant, B.F., Dawson, D.A., Stinson, F.S., Chou, P.S., Kay, W., Pickering, R., 2003. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. Drug Alcohol Depend. 71, 7–16.

Grant, B.F., Harford, T.C., Dawson, D.A., Chou, P.S., Pickering, R.P., 1995. The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability of alcohol and drug modules in a general population sample. Drug Alcohol Depend. 39, 37–44.

Grant, B.F., Stinson, F.S., Dawson, D.A., Chou, S.P., Ruan, W.J., Pickering, R.P., 2004. Co-occurrence of 12-month alcohol and drug use disorders and personality disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch. Gen. Psychiatry 61, 361–368.

Greenfield, T.K., Kerr, W.C., 2008. Alcohol measurement methodology in epidemiology: recent advances and opportunities. Addiction 103, 1082–1099.

Harford, T.C., Yi, H.Y., Faden, V.B., Chen, C.M., 2009. The dimensionality of DSM-IV alcohol use disorders among adolescent and adult drinkers and symptom patterns by age, gender, and race/ethnicity. Alcohol. Clin. Exp. Res. 33, 868–878.

Hasin, D., Aharonovich, E., Liu, X., Mamman, Z., Matseoane, K., Carr And, L.G., Li, T.K., 2002a. Alcohol dependence symptoms and alcohol dehydrogenase 2 polymorphism: Israeli Ashkenazis, Sephardics, and recent Russian immigrants. Alcohol. Clin. Exp. Res. 26, 1315–1321.

Hasin, D., Aharonovich, E., Liu, X., Mamman, Z., Matseoane, K., Carr, L., Li, T.K., 2002b. Alcohol and ADH2 in Israel: Ashkenazis, Sephardics, and recent Russian immigrants. Am. J. Psychiatry 159, 1432–1434.

Hasin, D., Carpenter, K.M., McCloud, S., Smith, M., Grant, B.F., 1997a. The alcohol use disorder and associated disabilities interview schedule (AUDADIS): reliability of alcohol and drug modules in a clinical sample. Drug Alcohol Depend. 44, 133–141.

Hasin, D., Grant, B.F., Cottler, L., Blaine, J., Towle, L., Ustun, B., Sartorius, N., 1997b. Nosological comparisons of alcohol and drug diagnoses: a multisite, multi-instrument international study. Drug Alcohol Depend. 47, 217–226.

Hasin, D., Paykin, A., 1999. Alcohol dependence and abuse diagnoses: concurrent validity in a nationally representative sample. Alcohol. Clin. Exp. Res. 23, 144–150.

Hasin, D., Rahav, G., Meydan, J., Neumark, Y., 1998. The drinking of earlier and more recent Russian immigrants to Israel: comparison to other Israelis. J. Subst. Abuse 10, 341–353.

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

Hasin, D.S., Van, R.R., McCloud, S., Endicott, J., 1997c. Differentiating DSM-IV alcohol dependence and abuse by course: community heavy drinkers. J. Subst. Abuse 9, 127–135.

Hasin, D.S., Stinson, F.S., Ogburn, E., Grant, B.F., 2007. Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: results from the National Epidemiologic survey on alcohol and related conditions. Arch. Gen. Psychiatry 64, 830–842.

Hu, L., Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct. Equat. Model. 26, 1–55.

Kahler, C.W., Strong, D.R., 2006. A Rasch model analysis of DSM-IV Alcohol abuse and dependence items in the National Epidemiologic Survey on Alcohol and Related Conditions. Alcohol. Clin. Exp. Res. 30, 1165–1175.

Keyes, K.M., Geier, T., Grant, B.F., Hasin, D.S., 2009. Influence of a drinking quantity and frequency measure on the prevalence and demographic correlates of DSM-IV alcohol dependence. Alcohol. Clin. Exp. Res. 33, 761–771.

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

Langenbucher, J.W., Labouvie, E., Martin, C.S., Sanjuan, P.M., Bavly, L., Kirisci, L., Chung, T., 2004. An application of item response theory analysis to alcohol, cannabis, and cocaine criteria in DSM-IV. J. Abnorm. Psychol. 113, 72–80.

Levav, I., Kohn, R., Golding, J.M., Weissman, M.M., 1997. Vulnerability of Jews to affective disorders. Am. J. Psychiatry 154, 941–947.

Li, T.K., Hewitt, B.G., Grant, B.F., 2007a. Is there a future for quantifying drinking in the diagnosis, treatment, and prevention of alcohol use disorders? Alcohol Alcohol. 42, 57–63.

Li, T.K., Hewitt, B.G., Grant, B.F., 2007b. The alcohol dependence syndrome, 30 years later: a commentary. the 2006 H. David Archibald lecture. Addiction 102, 1522–1530.

Lord, F.M., Novick, M.R., 1968. Statistical Theories of Mental Test Scores. Addison-Wesley, Reading, MA.

Martin, C.S., Chung, T., Kirisci, L., Langenbucher, J.W., 2006. Item response theory analysis of diagnostic criteria for alcohol and cannabis use disorders in adolescents: implications for DSM-V. J. Abnorm. Psychol. 115, 807–814.

Martin, C.S., Chung, T., Langenbucher, J.W., 2008. How should we revise diagnostic criteria for substance use disorders in the DSM-V? J. Abnorm. Psychol. 117, 561–575.

Muthén, L.K., Muthén, B.O., Chi-Square Difference Testing Using the Satorra-Bentler Scaled Chi-Square. http://www.statmodel.com/chidiff.shtml. Accessed on 27-12-2009.

Muthén, B.O., Muthén, L.K., 2008. Mplus: Statisitcal Analysis with Latent Variables (version 5.1). Muthén & Muthén Inc, Los Angeles, CA.

Muthén, L.K., Muthén, B.O., 1998–2007. Mplus User's Guide, fifth ed. Muthén & Muthén, Los Angeles, CA.

Neumark, Y.D., Rahav, G., Teichamn, M., Hasin, D., 2001. Alcohol drinking patterns among Jewish and Arab men and women in Israel. J. Stud. Alcohol 62, 443–447.

Norman, G.R., Streiner, D.L., 2008. Biostatistics: The Bare Essentials, third ed. BC Decker Inc, Hamilton, Ontario.

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

Pull, C.B., Saunders, J.B., Mavreas, V., Cottler, L.B., Grant, B.F., Hasin, D.S., Blaine, J., Mager, D., Ustun, B.T., 1997. Concordance between ICD-10 alcohol and drug use disorder criteria and diagnoses as measured by the AUDADIS-ADR, CIDI and SCAN: results of a cross-national study. Drug Alcohol Depend. 47, 207–216.

Rahav, G., Hasin, D., Paykin, A., 1999. Drinking patterns of recent Russian immigrants and other Israelis: 1995 national survey results. Am. J. Public Health 89, 1212–1216.

Saha, T.D., Chou, P.S., Grant, B.F., 2006. Toward an alcohol use disorder continuum using item response theory: results from the national epidemiologic survey on Alcohol and Related Conditions. Psychol. Med. 36, 931–941.

Saha, T.D., Stinson, F.S., Grant, B.F., 2007. The role of alcohol consumption in future classifications of alcohol use disorders. Drug Alcohol Depend. 89, 82–92.

SAS Institute Inc., 2008. SAS version 9.1.3. SAS Institute Inc., Cary, NC.

Spivak, B., Frisch, A., Maman, Z., Aharonovich, E., Alderson, D., Carr, L.G., Weizman, A., Hasin, D., 2007. Effect of ADH1B genotype on alcohol consumption in young Israeli Jews. Alcohol. Clin. Exp. Res. 31, 1297–1301.

Strong, D.R., Kahler, C.W., Colby, S.M., Griesler, P.C., Kandel, D., 2009. Linking measures of adolescent nicotine dependence to a common latent continuum. Drug Alcohol Depend. 99, 296–308.

Thissen, D., 2001. Irtlrdif v.20b: software for the computation of the statistics involved in item response theory likelihood ratio tests for differential item functioning. University of North Carolina, Chapel Hill, NC, unpublished manuscript.

Thissen, D., Steinberg, L., Wainer, H., 1988. Use of item response theory in the study of group differences in trace lines. In: Wainer, H., Braun, H. (Eds.), Test Validity. Erlbaum, Hillsdale, NJ, pp. 147–169.

Thissen, D., Steinberg, L., Wainer, H., 1993. Detection of differential item functioning using the parameters of item response models. In: Holland, P.W., Wainer, H. (Eds.), Differential Item Functioning. Erlbaum, Hillsdale, NJ, pp. 67–113.

Vrasti, R., Grant, B.F., Chatterji, S., Ustun, B.T., Mager, D., Olteanu, I., Badoi, M., 1998. Reliability of the Romanian version of the alcohol module of the WHO Alcohol Use Disorder and Associated Disabilities: Interview Schedule—alcohol/drug-revised. Eur. Addict. Res. 4, 144–149.

World Health Organization. Global Information System on Alcohol and Health. http://www.who.int/globalatlas/default.asp. Accessed on 27-12-2009.

Yeung, P.P., Greenwald, S., 1992. Jewish Americans and mental health: results of the NIMH Epidemiologic Catchment Area Study. Soc. Psychiatry Psychiatr. Epidemiol. 27, 292–297.