The Role of Alcohol Consumption in Future Classifications of Alcohol Use Disorders

Tulshi D. Saha, **Frederick S. Stinson**, and **Bridget F. Grant**
Division of Intramural Clinical and Biological Research, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, Maryland (Drs. Grant, Saha and Stinson)

Abstract

Background—Item response theory (IRT) was used to determine whether DSM-IV alcohol abuse and dependence and consumption criteria were arrayed along a continuum of severity.

Methods—Data came from a large, nationally representative sample of the U.S. adult population.

Results—DSM-IV alcohol abuse and dependence criteria formed a continuum of alcohol use disorder severity along with the drinking 5+/4+ at least once a week in the past year criterion. Criteria were invariant across sex, race-ethnicity, and age subgroups.

Conclusion—The drinking 5+/4+ high-risk drinking pattern was identified as a suitable criterion for future classifications of DSM-IV alcohol use disorder. Some dependence criteria were among the least severe criteria, and some abuse criteria were among the most severe, findings that question the validity of DSM-IV abuse and dependence categories as distinct entities and that do not support the assumption of abuse as prodromal to dependence. Physical dependence and addiction were identified as defining elements of the continuum. Further research examining their dimensional properties and relationships to high-risk drinking patterns appears warranted. An approach highlighting a more important role of consumption in future classifications of alcohol use disorder defined broadly to encompass all alcohol-related harm, including addiction and physical dependence is discussed.

Keywords

Alcohol use disorder; IRT analysis; addiction; physical dependence; high-risk drinking patterns

1. Introduction

In the *Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition* (DSM-IV; American Psychiatric Association, 1994), both alcohol abuse and alcohol dependence are defined as maladaptive patterns of alcohol use, leading to clinically significant impairment or distress, as manifested by three or more diagnostic symptom criteria for dependence and one

Corresponding Author/Author to Whom Requests for Reprints Should be Addressed: Bridget F. Grant, Ph.D., Ph.D., Laboratory of Epidemiology and Biometry, Room 3077, Division of Intramural Clinical and Biological Research, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, M.S. 9304, 5635 Fishers Lane, Bethesda, Maryland 20892-9304, Phone: 301-443-7370 Fax 301-443-1400, E-mail bgrant@willco.niaaa.nih.gov.

criterion or more for abuse. Despite the central importance of maladaptive patterns of drinking in the diagnostic classification of DSM-IV alcohol use disorders, no specific pattern of drinking is defined, nor do maladaptive patterns of drinking appear as diagnostic criteria for either disorder.

Historically, the absence of alcohol consumption in the DSM-IV can be traced back to early work on the classification of alcoholism conducted by Guze, Goodwin, and their colleagues at Washington University in St. Louis. Interestingly, the initial criteria published by Guze et al. in 1962 included quantity/frequency as a diagnostic criterion, defined by at least one of the following: (1) drinking every day; (2) drinking the equivalent in alcohol content of over 54 ounces of whisky per week; or (3) being unable to answer questions concerning frequency or quantity, interpreted to mean evasiveness about the amount of alcohol consumed (Guze et al. 1962; Goodwin et al. 1971). By 1967, however, the St. Louis Research Group had abandoned the quantity/frequency criterion because its reliability and validity had proved difficult to establish (Guze et al. 1967).

In what appears to be the first validation study of an alcoholism classification used in psychiatry, we find the reasoning underlying the abandonment of the quantity/frequency criterion (Guze et al. 1963). In this study, 223 male criminals' responses to questions assessing alcoholism criteria were compared with those of their relatives. However, the criterion concerning quantity/frequency of drinking was omitted from the comparisons because many of the relatives were unable to provide the information.

The impact of this decision by the St. Louis Research Group was far reaching. Research diagnostic criteria appearing after 1971, including the Feighner criteria (Feighner et al. 1972), the Research Diagnostic Criteria (RDC; Spitzer et al. 1978), the *Diagnostic and Statistical Manual of Mental Disorders – Third Edition* (DSM-III; American Psychiatric Association, 1980), *DSM – Third Edition – Revised* (DSM-III-R; American Psychiatric Association, 1987), and the *DSM – Fourth Edition* (DSM-IV; American Psychiatric Association, 1994), as well as successive revisions of the *International Classification of Diseases* (ICD-8; World Health Organization, 1967; ICD-9; World Health Organization, 1977; ICD-10; World Health Organization, 1992), did not include consumption-based criteria for any alcohol use disorder.

Recently published findings suggest that alcohol consumption might play an important role in DSM-IV classifications of alcohol use disorders. The results of two independent item response theory (IRT) analyses (Kahler and Strong, 2006; Saha et al. 2006) of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC: Grant et al. 2003, 2004) and one conducted using data from a large Australian national survey (Proudfoot et al. 2006) were consistent in demonstrating: (1) support for a single latent dimension underlying DSM-IV alcohol abuse and dependence criteria and associated symptom items; (2) the scarcity of criteria tapping the less severe range of the continuum; and (3) low severity and high discrimination associated with the larger/longer dependence criterion. Saha et al. (2006) suggested that the larger/longer criterion might serve as a bridging criterion that links the less severe end of the alcohol use disorder continuum to the more severe end that is characterized by physical dependence and compulsive drinking. Since the larger/longer criterion may represent a high-risk drinking pattern, these researchers suggested that other drinking patterns that confer risk, such as exceeding national drinking guidelines, may be good candidates to represent the mild end of the alcohol use disorder continuum.

In light of the questionable reasoning underlying the abandonment of the consumption criterion in alcohol use disorder classifications in psychiatry, and these new IRT findings, this study examines the effect of including a drinking pattern criterion in the DSM-IV classification of

alcohol use disorders. The selection of drinking pattern criteria in this study reflects a growing consensus in the consumption field that exceeding 4 drinks for men and 3 drinks for women per day, or per drinking occasion, reflects an important high-risk drinking pattern. This variable, often combined with frequency of exceeding those daily guidelines, underlies one aspect of the new NIAAA national drinking guidelines (NIAAA, 2005) and the new NIAAA definition of binge drinking (NIAAA, 2005), that is, a pattern of drinking that brings blood alcohol concentration to 0.08 gram percent, which corresponds to consuming 5 or more drinks (men) or 4 or more drinks (women) in about two hours. The 5+/4+ measure reflects the level of ethanol intake at which impairment in psychomotor and cognitive functioning is significantly increased (Dawson et al. 1996; Hindmarch et al. 1991; Turner et al. 1981). Further, epidemiologic risk curve analyses show strong and rapid increases in the risks of accidental injuries (Cherpitel et al. 1995, 1996, 2006), deaths from external causes (Dawson, 2001), being a target of aggression or involved in an aggressive incident (Dawson, 1997; Graham and West, 2001; Room et al. 1995; Rossow, 1996), alcohol use disorders (Dawson et al. 2005a, 2005b, 2005c), and numerous adverse social, medical, occupational, and legal outcomes of drinking (Dawson, 2000; Greenfield, 2001; Midanik et al. 1996; Rehm et al. 2005; Russell et al. 2004; Wechsler and Nelson, 2006) as the frequency of consuming 5+/4+ drinks rises, even at lower levels of frequency.

Thus, the purpose of the present study was to determine whether the DSM-IV abuse and dependence criteria, along with the 5+/4+ drinking pattern criterion, defined an alcohol use disorder continuum, with drinking pattern criteria being able to tap the mildest end of the severity continuum as suggested by prior research. IRT analysis was used to examine whether these criteria could be represented by a single underlying latent dimension using a large (n=43,093), nationally representative survey of the U.S. general population, the NIAAA NESARC (Grant et al. 2003, 2004). Because risk of alcohol-related harm has been empirically demonstrated using several levels of frequency of drinking 5+/4+, three levels of frequency of exceeding these daily guidelines were examined: (1) at least once in the past year; (2) at least 12 times in the past year; and (3) at least once a week in the past year. In addition, the large sample size of the NESARC allowed for the examination of differential criterion function (DCF) versus invariance of the criteria across sex, age, and race-ethnic subgroups of the population, analyses infrequently conducted in past research due to small sample sizes.

2. Methods

2.1. Sample

The 2001–2002 NESARC is a representative sample of the United States conducted by NIAAA, as described elsewhere (Grant et al. 2003, 2004). The NESARC target population was the civilian population residing in households and group quarters, 18 years and older. Face-to-face interviews were conducted with 43,093 respondents, with a response rate of 81%. Blacks, Hispanics, and young adults (ages 18 – 24 years) were oversampled, with data adjusted for oversampling and household- and person-level nonresponse. The weighted data were then adjusted to represent the U.S. civilian population based on the 2000 census. The sample for this study was restricted to 20,846 respondents classified as regular current drinkers (i.e., those who drank at least 12 drinks in the past year).

2.2. DSM-IV Alcohol Abuse and Dependence Criteria

The NIAAA Alcohol Use and Disorders and Associated Disabilities Interview Schedule – DSM-IV version was designed to generate DSM-IV diagnoses of alcohol abuse and alcohol dependence (Grant et al. 2001). These diagnoses were based on 36 symptom items that formed the basis of the four criteria for alcohol abuse and seven criteria for alcohol dependence appearing in the DSM-IV. The symptom items underlying the criteria appear on the NIAAA

web site and have been published elsewhere (Grant et al. 2006). Following DSM-IV, diagnoses of current alcohol abuse required one of the four criteria to be met during the 12 months prior to the interview, while dependence diagnoses required at least 3 associated diagnostic criteria to be met during that time.

As reported in detail elsewhere, the reliability and validity were good to excellent for alcohol use disorder diagnoses (Muthén et al. 1993; Grant et al. 1995, 2003, 2004; Chatterji et al. 1997; Cottler et al. 1997; Hasin et al. 1997; 2003; Pull et al. 1997; Ustun et al. 1997; Vrasti et al. 1997; Canino et al. 1999; Hasin and Paykin, 1999; Nelson et al. 1999; Compton et al. 2004). Reliability (Chatterji et al. 1997) and validity (Cottler et al. 1997) for alcohol abuse and dependence criteria were fair to good as assessed in clinical reappraisal studies conducted by psychiatrists using a semi-structured diagnostic interview. Intraclass correlations of alcohol abuse and dependence criteria derived from a test-retest study of the general population (Grant et al. 1995, 2003) were excellent (intraclass correlations=0.71–0.75) as were the kappa values (κ=0.61–0.74).

2.3. Statistical Analysis

From several IRT models, we selected the two-parameter logistic model (Birnbaum, 1968; Lord and Novick, 1968), an extension of the Rasch or one-parameter logistic model (Rasch, 1960), to define the relationship between the observed responses to the criteria and the underlying unobserved latent trait or construct (alcohol use disorder severity). This IRT model, generated using the BILOG-MG statistical program (Scientific Software International, 2003), yields marginal maximum likelihood estimates (Bock and Aitkin, 1981; Harwell et al. 1988) of two parameters: the b or threshold parameter and the a or discrimination parameter. The a parameter measures the ability of a criterion to discriminate people who are higher on the continuum and those who are lower on the continuum. This parameter describes how strongly the criterion is related to the underlying trait or construct. The larger the a parameter (i.e., the slope at its steepest point), the greater the discrimination of a criterion. The b or threshold parameter measures the severity of a criterion; criteria with high thresholds are endorsed less frequently and are more severe.

The a and b parameters were plotted graphically as criterion response curves (CRCs). In these plots the b parameter represents the criterion's location along the latent continuum (located on the horizontal axis). The b parameter (severity) is the point on the latent continuum where there is a 50% chance of the criterion being present. The b parameter shifts the CRC from left to right as the criterion becomes more severe. The a or discrimination parameter indicates how steep the slope of the CRC is at its steepest point. Criterion severity is an indication of the degree of alcohol problem severity that is needed for a particular criterion to be present.

To be clinically useful, criterion items should be shown to be invariant across important subgroups of the population, defined in terms of sex, race-ethnicity and age. The presence of DCF would suggest differences may exist in the expression of certain alcohol use disorder criteria which would contribute bias in the ordering of individuals along the underlying continuum. To determine whether abuse, dependence, or 5+/4+ criteria displayed DCF, we statistically compared a and b parameters for each criterion across groups defined by sex (men as the referent category), age (18 – 29 as the referent category, 30 – 44, and 45+ years) and race-ethnicity (Black as the referent category, White and Hispanic) using IRT methodology. These DCF analyses were conducted using the PARSCALE program (Scientific Software International, 2003).

Criteria that demonstrate DCF need not reflect bias or variance across subgroups if the DCF occurs in opposing directions (e.g., some criteria result in greater discrimination or severity among men while others demonstrate the opposite effect) (Cooke et al. 2001; Bolt et al.

2004). Whether DCF actually reflects invariance across subgroups can be determined if the observed criterion-level DCFs cancel out at the total test (scale) score level. To accomplish this, we plotted the expected raw scores by the severity of the alcohol use disorder continuum for age, sex and race-ethnic groups – plots referred to as the test response curves (TRCs). If the TRCs for subgroups (e.g., for men and women) do not substantially differ we can conclude that the significant criterion-level DCFs (if found) cancel out when considered at the total scale level and that for any latent trait value, men and women have identical expected raw scores. If, however, the TRCs do substantially differ by sex, age or race-ethnicity, criteria demonstrating DCF can be assumed to be biased, lacking invariance across important subgroups of the population.

As a preliminary to the above-referenced IRT analyses, we conducted exploratory factor analyses using Mplus software (Muthén and Muthén, 2004) in order to test the IRT assumption of unidimensionality. Although Lord (1952) initially emphasized the importance of the unidimensionality assumption, more recent evidence has shown that with sufficient sample sizes and the dominance of one dimension, IRT parameter estimates are stable and accurately represent the underlying data (Reckase, 1979; Dragow and Parsons, 1983; Harrison, 1986; Hambleton, 1989; Drasgow and Hulin, 1990). However, the parameters of the IRT model are most interpretable when the criteria reflect a single unitary dimension (Stout, 1987; Downing, 2003; Bolt et al. 2004; Krueger et al. 2004).

3. Results

3.1. Prevalence and Factor Analysis

The prevalences of DSM-IV abuse and dependence criteria ranged from 1.15% for activities given up and legal problems to 13.84% for larger/longer (Table 1). The rates of exceeding the daily limit at least once, at least 12 times, and at least once a week during the past year were 49.15%, 30.18%, and 18.54%, respectively.

Exploratory factor analyses (Muthén and Muthén, 2004) were performed to assess the IRT unidimensionality assumption separately for models including each frequency of the 5+/4+ drinking pattern criterion (Table 1). Several criteria were used to examine the unidimensionality of the data for each model, including: (1) the eigenvalues of the correlation matrix associated with the one-factor, two-factor and three-factor models; and (2) the amount of variance explained by the first factor. The eigenvalues for the one-, two- and three-factor solutions were by far greater for the first factor (8.098–8.133) than for the second factor (0.712–0.724) or the third factor (0.536–0.578), regardless of which consumption criterion model was examined. The ratio of the first-to-second eigenvalues, a common criterion for assessment of unidimensionality, was more than 11 for all groups. Further, Reckase (1979) concluded that an eigenvalue that accounted for at least 20% of the total variance is needed for stable IRT parameters. Thus, the one-factor model was selected over the two- or three-factor model based on the high proportion of variance explained by it (68%).

Exploratory factor analysis conducted in a confirmatory factor analytic framework using Mplus (Muthén and Muthén, 2004) also produces both goodness of fit statistics and chi-square statistics that can be used to compare once-, two- and three-factor solutions. For each model, the chi-square statistic was highly significant. However, because the chi-square statistic is highly sensitive to large sample sizes (as in the case of ours) and may overstate the lack of fit of a structural model (Bollen, 1989), this test statistic was not used to compare factor solutions. However, a number of additional fit indices have been developed that attempt to account for this problem associated with the chi-square statistic. Hu and Bentler (1999) provided a test of the "rules of thumb" cutoffs for the most commonly used fit indices. They advocated a two-index strategy to assess the adequacy of fit of structural models. Hu and Bentler (1999) suggest

that a cutoff of 0.95 or above on either the Tucker Lewis Index (TLI; Tucker and Lewis, 1973) or the Comparative Fit Index (CFI; Bentler, 1990) combined with the root mean squared error of approximation (RMSEA) "close to 0.06" provides a good combination of fit indices to conclude a good fit of a factor model. Consistent with the recommendations of Hu and Bentler (1999), the one-factor solution demonstrated a good fit to the observed data (Table 1).

3.2. Criterion Response Curves

Table 1 also presents the IRT parameters for the two-parameter logistic models. The two-parameter logistic models were also found to provide a slightly better fit to the data (Bayesian Information Functions – BIC = 97610.48 – 105471.87) relative to the one-parameter Rasch models (97733.83 – 105476.22) for each consumption measure. The associated BIC was somewhat lower for the highest frequency of the 5+/4+ drinking pattern criterion, indicating slightly improved fit.

The CRCs associated with each model are shown in Figures 1–3. As expected, severity increased as the frequency of consuming 5+/4+ drinks increased, from −0.04 for exceeding the daily limits once in the past year to 1.22 for exceeding the limits at least once a week in the past year. However, all drinking pattern criteria fell along the less severe end of the alcohol use disorder continuum. In all three models, severity was greatest for the legal problems, activities given up, and neglect of roles criteria, and lowest for exceeding the daily consumption limits, larger/longer, quit/control, and hazardous use criteria, with the remaining criteria representing intermediate severity levels. There was less variability observed in discrimination, with the larger/longer, time spent, and physical/psychological problems criteria associated with greatest discrimination, and legal problems, hazardous use, and quit/control associated with lowest discrimination.

3.3. Differential Criterion Functioning

There were relatively few criteria that exhibited DCF at the criterion-level by sex and race-ethnicity. Significant differences ($p < 0.05$) in severity and discrimination were observed for 2 to 3, and 5 to 6, of the criteria, respectively, between men and women, in each model. The corresponding number of criterion differences observed in severity and discrimination between Blacks and Whites were 3 to 4 and 3 to 4, respectively. Discrimination parameters for each criterion were not significantly different between Blacks and Hispanics, whereas severity parameters were only significantly different for the consumption criterion in two of the three models. In contrast, DCF was much greater for age at the criterion- level in all three models. The greatest DCF was exhibited in severity and discrimination between 18–29- and 30–44-year-olds, with 4 to 8 and 2 to 10 criteria, respectively, differing significantly between these groups. Fewer significant differences were observed in terms of severity (2 to 5 criteria) and discrimination (6 to 8 criteria) between 18–29 year-olds and respondents 45 years and older.

DCF associated with severity and discrimination parameters by sex, race-ethnicity, and age, occurred in both directions, regardless of which drinking pattern criterion was entered into the model. Consistent with this observation, TRCs were virtually identical for each sex, race-ethnic group, and age group. These curves were also identical for each of the three models and for illustrative purposes are shown for the model that included the 5+/4+ drinking pattern at least once a week in the past year (Figures 4A–4C). Although some abuse, dependence, and consumption criteria did exhibit DCF at the criterion-level, there was no evidence of DCF between subgroups at the total score level.

4. Discussion

The results of this study showed that DSM-IV alcohol abuse and dependence criteria, along with a 5+/4+ drinking pattern criterion assessed at three different levels of frequency, were arrayed along a continuum of severity. These results question the validity of the DSM-IV abuse and dependence categories as distinct entities. The criterion information curves identified some dependence criteria (e.g., larger/longer and quit/control) as being among the mildest, while some abuse criteria (e.g., neglect of roles) were included in the more severe range of the continuum, findings that do not support the frequent assumption that abuse is prodromal to dependence. Further, there was no evidence of differential criterion functioning, indicating criterion invariance across sex, age, and race-ethnic subgroups of the population.

Consistent with prior research (Langenbucher et al. 2004; Proudfoot et al. 2006; Saha et al. 2006), the abuse and dependence criteria tapped the more severe range of the underlying continuum. Also consistent with prior research (Krueger et al. 2004; Langenbucher et al. 2004; Proudfoot et al. 2006; Kahler and Strong, 2006; Saha et al. 2006), the larger/longer dependence criterion was shown to have excellent discrimination and low severity. Taken together, these results suggest that the larger/longer criterion may represent a high-risk drinking pattern that serves as a bridging criterion linking the less severe end of the alcohol use disorder continuum with the more severe end. If this is the case, as previously mentioned, then we would expect other high-risk drinking patterns to be good candidates to represent the milder end of the continuum. This is precisely what was found. Each of the three 5+/4+ drinking pattern criteria tapped into the mildest range of the alcohol use disorder continuum, at lower levels of severity than those captured by any of the abuse or dependence criteria.

The severity of the drinking pattern criterion increased as a function of frequency of consuming 5+/4+, from −0.04 at least once in the past year to 0.66 for at least 12 times to 1.22 for exceeding the guidelines at least once a week. The best-fitting model included drinking 5+/4+ at least once a week during the past year. In addition to model fit, the selection of a specific frequency of the 5+/4+ drinking pattern criterion should be based on the measure's ability to serve the clinical utility of a diagnostic classification. Based on model fit and this selection criterion, we chose consumption of 5+/4+ drinks at least once a week in the past year as the optimal candidate for inclusion in the DSM-V revision of the alcohol use disorder classification. This criterion taps a milder end of the alcohol use disorder continuum than any existing abuse or dependence criterion and represents severity at the intermediate to severe range of the observed continuum, thereby preserving the utility of the classification to identify clinically significant cases of alcohol use disorder in need of treatment. The drinking 5+/4+ at least once a year and drinking 5+/4+ at least 12 times in the past year criteria were too mild in terms of severity to serve this major purpose of the DSM-V classification.

The relative ordering of the severity of abuse and dependence criteria in the present study are generally consistent with, but differ in some ways from, prior research. Neglect of roles, activities given up, and legal problems were among the most severe criteria, a result consistent with previous studies whether conducted at the item (Kahler and Strong, 2006) or criterion (Langenbucher et al. 2004; Proudfoot et al. 2006;) level. Psychological/physical problems, social/interpersonal problems, and time spent criteria were also in the severe range of the observed continuum in this study, consistent with items tapping similar constructs in the studies by Krueger et al. (2004) and Kahler and Strong (2006), but at variance with findings reported by Langenbucher et al. (2004) that showed psychological/physical problems to be the least severe criterion, with the remaining criteria overlapping in the middle range of the observed severity continuum. Similar to prior research in this area, the quit/control, larger/longer, and hazardous use criteria were among the least severe (Krueger et al. 2004; Kahler and Strong, 2006).

The withdrawal dependence criterion fell within the middle range of the observed continuum, a finding consistent with IRT analyses of withdrawal items conducted by Krueger et al. (2004) and Kahler and Strong (2006). In those two studies, fits and seizures, visual distortions, and hallucinations were among the most severe; depression, sweating, anxiety, and insomnia were intermediate in terms of severity. Interestingly, vomiting/sick to stomach, another DSM-IV withdrawal category symptom, was the least severe in the study by Kahler and Strong (2006) and the present study, but among the most severe in the study by Krueger et al. (2004). Thus, given the range of severity of DSM-IV withdrawal items, it is not surprising that the aggregate withdrawal criterion in this study fell in the middle range of severity.

That DSM-IV dependence criteria relate to a single underlying dimension of alcohol use disorder severity is not surprising since the dependence criteria operationalize either physical dependence (i.e., tolerance and withdrawal) or loss of control or compulsive drinking (i.e., addiction as defined in the neuroscience literature), two closely related concepts. In this regard, it is important to note that three of the four abuse criteria operationalize a broader definition of compulsive drinking, that is, drinking despite social/interpersonal problems or recurrent drinking that results in neglect of role responsibilities or legal problems. In contrast, the fourth abuse criterion, recurrent drinking in situations likely to cause physical harm to the drinker and/or others, is an indicator of potential harm or a risk factor, not loss of control or physical dependence, both of which constitute consequences of pathological consumption. Thus, at the conceptual level, the utility of this abuse criterion warrants further investigation despite its strong relationship to the observed underlying continuum. Taken together, these results argue strongly for the abandonment of the alcohol abuse category and the inclusion of at least three of its criteria within a single alcohol use disorder category. Within the context of future research, drinking despite actual alcohol-related harm resulting from use in hazardous situations, such as unintentional injury as a result of driving after drinking too much, could be examined as an additional indicator of compulsive drinking suitable for the alcohol use disorder category.

Criteria representing physical dependence and loss of control/compulsion for the most part defined the observed alcohol use disorder severity continuum, a result that is not entirely consistent with research on the neurobiology of addiction. Human imaging studies and animal self-administration models have emphasized that physical dependence (tolerance and withdrawal) is not a necessary condition for the initiation of compulsive drinking (addiction), though it may play a role in the maintenance of addiction, and that neuroadaptations at the molecular and cellular levels underlying physical dependence on alcohol and addiction may be quite different (Koob et al., 1998; Charney and Nestler, 2004). Consistent with these general findings from neurobiology, the DSM-III-R Substance Use Disorder Work Group did unsuccessfully propose the need for a psychosubstance neuroadaptation category for individuals who demonstrated physical dependence without the behavioral or cognitive elements of loss of control and compulsion (Rounsaville et al. 1986). In their view, individuals with the neuroadaptation category had a disorder, but not a mental disorder, and the category would best be placed on Axis III (general medical conditions).

Members of the DSM-III-R Work Group and the Chairperson of the DSM-V Work Group (O'Brien et al. 2006) additionally recognized the potentially stigmatizing aspects of a dependence diagnosis, a diagnosis that should be reserved for those who additionally display indicators of loss of control and compulsion. That is, physical dependence can occur with any drug that affects the central nervous system, even if prescribed by a physician and taken as prescribed. In these situations, absent indicators of compulsive use or loss of control, a separate diagnosis other than dependence may be warranted. In view of this, it would make nosologic sense to examine the dimensional properties of physical dependence and addiction separately in future research, along with consumption patterns that relate to them that may well differ.

This research will importantly help address the larger issue as to what extent advances in neuroscience and other disciplines can inform clinical diagnosis of alcohol use disorders.

The IRT analyses importantly provided for a comparison of several consumption criteria and an empirical basis for the selection of a consumption criterion that would be clinically meaningful. In doing so, one important unspecified aspect of DSM-IV definitions of alcohol use disorder, that is a maladaptive pattern of alcohol use, has been explicitly defined. The objective nature of the high-risk drinking pattern criterion clearly defines its clinical utility in the assessment of alcohol use disorder. Moreover, the high-risk IRT results presented herein can be used to create a continuous or dimensional scale of alcohol use disorder by weighting each of the diagnostic criteria by their associated severity weights or severity and discrimination weights. Similar scales can be constructed for physical dependence and addiction if further research supports this distinction. In both cases, the scales would be directly linked to the DSM-IV classification the continuum has helped to define. However, the results of this study beg the question of how to establish a clinically relevant cutpoint or threshold for the categorical diagnosis of alcohol use disorder, a central nosologic issue in future research concerning all psychiatric disorders appearing in the nomenclature.

This study identified the frequency of drinking 5+/4+ drinks that could serve to define, in part, the alcohol use disorder continuum characterized by physical dependence and addiction. However, alcohol consumption may have a much larger role to play in developing a future unitary classification of alcohol disorders that will encompass all actual alcohol-related harm, including physical, psychological, economic, social, and legal harm, in addition to physical dependence and addiction. In this classification, each type of alcohol-related harm could be linked to various cutpoints along a single continuous measure of high-risk drinking. This classification, based on the quantification of a single high-risk drinking pattern or trait with cutpoints associated with specific alcohol-related harms, could resemble classifications of other complex diseases, such as blood pressure for hypertension, fasting plasma glucose for Type II diabetes mellitus, and body mass index for obesity.

Acknowledgments

References

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorder. Vol. 3. Washington, D.C: American Psychiatric Association; 1980.

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Vol. 3. Washington, D.C: American Psychiatric Association; 1987.

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Vol. 4. Washington, D.C: American Psychiatric Association; 1994.

Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990;107:203–246.

Birnbaum, A. Some latent trait models. In: Lord, FM.; Norvick, MR., editors. Statistical Theory of Mental Test Scores. Addison-Wesley; Reading, MA: 1968. p. 397-472.

Bock RD, Aitkin M. Marginal maximum likelihood estimation of item parameters: application of an EM algorithm. Psychometrika 1981;46:443–445.

Bollen, KA. Structural Equations with Latent Variables. Wiley; New York, NY: 1989.

Bolt DM, Hare RD, Vitale JE, Newman JP. A multigroup item response theory analysis of the Psychopathy Checklist-revised. Psychol Assess 2004;16:155–168. [PubMed: 15222812]

Canino GJ, Bravo M, Ramírez R, Febo V, Fernandez R, Hasin DS. The Spanish Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability and concordance with clinical diagnosis in a Hispanic population. J Stud Alcohol 1999;60:790–799. [PubMed: 10606491]

Charney, DS.; Nestler, EJ., editors. Neurobiology of Mental Illness. Oxford University Press; Oxford, England: 2004.

Chatterji S, Saunders JB, Vrasti R, Grant BF, Hasin D, Mager D. 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 1997;47:171–185. [PubMed: 9306043]

Cherpitel CJ, Tan T, Midanik L, Caetano R, Greenfield T. Alcohol and non-fatal injury in the U.S. general population: a risk function analysis. Accid Anal Prev 1995;27:651–661. [PubMed: 8579696]

Cherpitel CJ, Bond J, Ye Y. Alcohol and injury: a risk function analysis from the Emergency Room Collaborative Alcohol Analysis Project (ERCAAP). Eur Addict Res 2006;12:42–52. [PubMed: 16352902]

Compton WM, Grant BF, Colliver JD, Glantz MD, Stinson FS. Prevalence of marijuana use disorders in the United States: 1991–1992 and 2001–2002. JAMA 2004;291:2114–2121. [PubMed: 15126440]

Cooke DJ, Kosser DS, Michie C. Psychopathy and ethnicity: structural, item and test generalizability of the Psychopathy Checklist – Revised (PCL-R) in Caucasian and African American participants. Psychol Assess 2001;13:531–542. [PubMed: 11793896]

Cottler LB, Grant BF, Blaine J, Mavreas V, Pull C, Hasin DS, Compton WM, Rubio-Stripec M, Mager D. Concordance of DSM-IV alcohol and drug use disorder criteria and diagnoses as measured by AUDADIS-ADR, CIDI and SCAN. Drug Alcohol Depend 1997;47:195–205. [PubMed: 9306045]

Dawson DA. Alcohol, drugs, fighting and suicide attempt/ideation. Addict Res 1997;5:451–472.

Dawson DA. Alcohol and mortality from all causes. J Stud Alcohol 2001;62:790–797. [PubMed: 11838916]

Dawson DA, Grant BF, Li TK. Quantifying the risks associated with exceeding recommended drinking limits. Alcohol Clin Exp Res 2005a;29:902–908. [PubMed: 15897737]

Dawson DA, Grant BF, Stinson FS, Zhou Y. Effectiveness of the derived alcohol use disorders identification test (AUDIT-C) in screening for alcohol use disorders and risk drinking in the U.S. general population. Alcohol Clin Exp Res 2005b;29:844–854. [PubMed: 15897730]

Downing M. Item response theory: applications of modern test theory in medical education. Med Educ 2003;37:739–745. [PubMed: 12945568]

Drasgow, F.; Hulin, CL. Item response theory. In: Dunnette, MD.; Hough, LM., editors. Handbook of Industrial and Organizational Psychology. Vol. 2. Vol. 2. Consulting Psychologists Press; Palo Alto, CA: 1990. p. 577-636.

Drasgow F, Parsons CK. Application of unidimensional item response theory models to multidimensional data. Appl Psychol Meas 1983;7:189–199.

Feighner JP, Robins E, Guze SB, Woodruff RA, Winokur G, Munoz R. Diagnostic criteria for use in psychiatric research. Arc Gen Psychiatry 1972;26:57–63.

Goodwin DW, Crane JB, Guze SB. Felons who drink: an 8-year follow-up. Q J Stud Alcohol 1971;32:136–147. [PubMed: 5546040]

Graham, K.; West, P. Alcohol and crime. In: Heather, N.; Peters, TJ.; Stockwell, T., editors. International Handbook of Alcohol Dependence and Problems. John Wiley and Sons; Chichester, England: 2001. p. 439-470.

Grant BF, Harford TC, Dawson DA, Chou SP, Pickering RP. The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability of alcohol and drug modules in a general population sample. Drug Alcohol Depend 1995;39:37–44. [PubMed: 7587973]

Grant BF, Harford TC, Muthén BO, Yi H, Hasin DS, Stinson FS. DSM-IV alcohol dependence and abuse: further concurrent validity in the general population. Drug Alcohol Depend. 2006[Epub ahead of print]

Grant, BF.; Moore, TC.; Shepard, J.; Kaplan, K. Source and Accuracy Statement, Wave 1 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). National Institute on Alcohol Abuse and Alcoholism; Bethesda, MD: 2003.

Grant BF, Stinson FS, Dawson DA, Chou SP, Dufour MC, Compton WM, Pickering RP, Kaplan K. Prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arc Gen Psychiatry 2004;61:807–816.

Greenfield, TK. Individual risk of alcohol-related diseases and problems. In: Heather, N.; Peters, TJ.; Stockwell, T., editors. International Handbook of Alcohol Dependence and Problems. John Wiley and Sons; Chichester, England: 2001. p. 413-437.

Guze SB, Tuason VB, Gatfield PD, Stewart MA, Picken B. Psychiatric illness and crime with particular reference to alcoholism: a study of 223 criminals. J Nerv Ment Dis 1962;134:512–521. [PubMed: 13903238]

Guze SB, Tuason VB, Stewart MA, Picken B. The drinking history: a comparison of reports by subjects and their relatives. Q J Stud Alcohol 1963;24:249–260. [PubMed: 13951521]

Guze SB, Wolfgram ED, McKinney JK, Cantwell DP. Psychiatric illness in families of convicted felons: a study of 519 first-degree relatives. Dis Nerv Syst 1967;28:651–659. [PubMed: 6051292]

Hambleton, RK. Principles and selected applications of item response theory. In: Linn, RL., editor. Educational Measurement. Vol. 3. American Council on Education, Macmillan Publishing; New York, NY: 1989. p. 147-200.

Harrison DA. Robustness of IRT parameter estimation to violations of the unidimensionality assumption. J Educ Stat 1986;11:91–115.

Harwell MR, Baker FB, Zwarts M. Item parameter estimation via marginal maximum likelihood and an EM algorithm: a didactic. J Educ Stat 1988;13:243–271.

Hasin, DS.; Muthén, B.; Grant, BF. The dimensionality of DSM-IV alcohol abuse and dependence: factor analysis in a clinical sample. In: Vrasti, R., editor. Alcoholism: New Research Perspectives. Hogrefe and Hubner; Munich, Germany: 1997. p. 36-52.

Hasin DS, Paykin A. Alcohol dependence and abuse diagnoses: concurrent validity in a nationally representative sample. Alcohol Clin Exp Res 1999;23:144–150. [PubMed: 10029216]

Hasin DS, Schuckit MA, Martin CS, Grant BF, Bucholz KK, Helzer JE. The validity of DSM-IV alcohol dependence: what do we know and what do we need to know? Alcohol Clin Exp Res 2003;27:244–252. [PubMed: 12605073]

Hindmarch I, Kerr JS, Sherwood N. The effects of alcohol and other drugs on psychomotor performance and cognitive function. Alcohol Alcohol 1991;26:71–79. [PubMed: 1854375]

Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling 1999;26:1–55.

Kahler CW, Strong DR. 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 2006;30:1165–1175. [PubMed: 16792564]

Koob GF, Sanna PP, Bloom FE. Neuroscience of addiction. Neuron 1998;21:467–476. [PubMed: 9768834]

Krueger RF, Nicol PE, Hicks BM, Markon KE, Patrick CJ, Iacono WG, Mague M. Using latent trait modeling to conceptualize an alcohol problems continuum. Psychol Assess 2004;16:107–119. [PubMed: 15222807]

Langenbucher JW, Labouvie E, Martin CS, Sanjuan PM, Bavly L, Kirisci L. An application of item response theory analysis to alcohol, cannabis and cocaine criteria in DSM-IV. J Abnorm Psychol 2004;113:72–80. [PubMed: 14992659]

Lord, FM. A Theory of Test Scores. Psychometric Monographs, No. 7. University of Chicago; Chicago, IL: 1952.

Lord, FM.; Novick, MR. Statistical Theories of Mental Test Scores. Addison-Wesley; Reading, MA: 1968.

Midanik LT, Tam T, Greenfield TK, Caetano R. Risk functions for alcohol-related problems in a 1988 U.S. national sample. Addiction 1996;91:1427–1437. [PubMed: 8917925]

Muthén BO. Factor analysis of alcohol abuse and dependence symptom items in the 1988 National Health Interview Survey. Addiction 1995;90:637–645. [PubMed: 7795500]

Muthén BO, Hasin D, Wisnicki KS. Factor analysis of ICD-10 symptom items in the 1988 National Health Interview Survey on Alcohol Dependence. Addiction 1993;88:1071–1077. [PubMed: 8401161]

Muthén, BO.; Muthén, LK. Mplus: Statistical Analysis with Latent Variables (Version 3.01). Los Angeles, California: Muthén & Muthén Inc; 2004.

National Institute on Alcohol Abuse and Alcoholism. Task Force Report on Binge Drinking. National Institute on Alcohol Abuse and Alcoholism; Rockville, MD: 2004.

National Institute on Alcohol Abuse and Alcoholism. Helping Patients Who Drink Too Much: A Clinician's Guide. National Institute on Alcohol Abuse and Alcoholism; Rockville, MD: 2005.

Nelson CB, Rehm J, Ustun TB, Grant BF, Chatterji S. Factor structure for DSM-IV substance disorder criteria endorsed by cannabis, cocaine and opioid users: results of the World Health Organization Reliability and Validity Study. Addiction 1999;94:843–855. [PubMed: 10665074]

O'Brien CP, Volkow N, Li TK. What's in a word? Addiction versus dependence in DSM-V. Am J Psychiatry 2006;163:764–765. [PubMed: 16648309]

Proudfoot H, Baillie AJ, Teesson M. The structure of alcohol dependence in the community. Drug Alcohol Depend 2006;81:21–26. [PubMed: 16005578]

Pull CB, Saunders JB, Mavreas V, Cottler LB, Grant BF, Hasin DS, Blaine J, Mager D, Ustun BT. 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 1997;47:207–216. [PubMed: 9306046]

Rasch, G. Probabilistic Models for Some Intelligence and Attainment Tests. Danish Institute for Educational Research; Copenhagen, Denmark: 1960.

Rehm J, Monga N, Adlaf E, Taylor B, Bondy SJ, Fallu JS. School matters: drinking dimensions and their effects on alcohol-related problems among Ontario school students. Alcohol Alcohol 2005;40:569–574. [PubMed: 16157609]

Reckase MD. Unifactor latent trait models applied to multifactor tests: results and implications. J Educ Stat 1979;4:207–230.

Room R, Bondy SJ, Ferris J. Risk of harm to oneself from drinking. Canada Addiction 1995;90:499–513.

Rossow I. Alcohol-related violence: the impact of drinking pattern and drinking context. Addiction 1996;91:1651–1661. [PubMed: 8972923]

Rounsaville BJ, Spitzer RL, Williams JBW. Proposed changes in DSM-III substance use disorders: description and rationale. Am J Psychiatry 1986;143:463–468. [PubMed: 3953889]

Russell M, Light JM, Gruenewald PJ. Alcohol consumption and problems. The relevance of drinking patterns. Alcohol Clin Exp Res 2004;28:921–930. [PubMed: 15201635]

Saha TD, Chou SP, Grant BF. Toward an alcohol use disorder continuum using item response theory: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Psychol Med 2006;36:931–941. [PubMed: 16563205]

Du Toit, M., editor. Scientific Software International. Item Response Theory (IRT) from SSI. Scientific Software International; Lincolnwood, IL: 2003.

Spitzer RL, Endicott J, Robins E. Research diagnostic criteria: rationale and reliability. Arc Gen Psychatry 1978;35:773–782.

Stout W. A new item response theory modeling approach with application of unidimensional assessment and ability estimation. Psychometrika 1987;55:293–326.

Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. Psychometrika 1973;38:1–10.

Turner C. How much alcohol is in a standard drink? An analysis of 125 studies. Br J Addict 1990;85:1171–1175. [PubMed: 2224197]

Ustun B, Compton WM, Mager D, Babor T, Baiyewu O, Chatterji S, Cottler LB, Cogus A, Mavreas V, Peters L, Pull D, Saunders J, Smeets R, Stipec MR, Vrasti R, Hasin D, Room R, Van De Brink W, Regier D, Baline J, Grant BF, Sartorius N. WHO study on the reliability and validity of the alcohol and drug use disorder instruments: overview of methods and results. Drug Alcohol Depend 1997;47:161–169. [PubMed: 9306042]

Vrasti R, Grant BF, Chatterji S, Ustun BT, Mager D, Olteanu I, Nodoi M. The reliability of the Romanian version of the alcohol module of the WHO Alcohol Use Disorder and Associated Disabilities Interview Schedule-Alcohol/Drug-Revised (AUDADIS-ADR). Eur Addict Res 1997;4:144–149. [PubMed: 9852366]

Wechsler H, Nelson TF. Relationship between levels of consumption and harms in assessing drink cut points for alcohol research: commentary on "many college freshmen drinking at levels far beyond the binge threshold: by White et al. Alcohol Clin Exp Res 2006;30:922–937. [PubMed: 16737449]

World Health Organization. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death. Vol. 8. World Health Organization; Geneva, Switzerland: 1967.

World Health Organization. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death. World Health Organization; Geneva, Switzerland: 1977. Ninth Revision

World Health Organization. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death. World Health Organization; Geneva, Switzerland: 1992. Tenth Revision
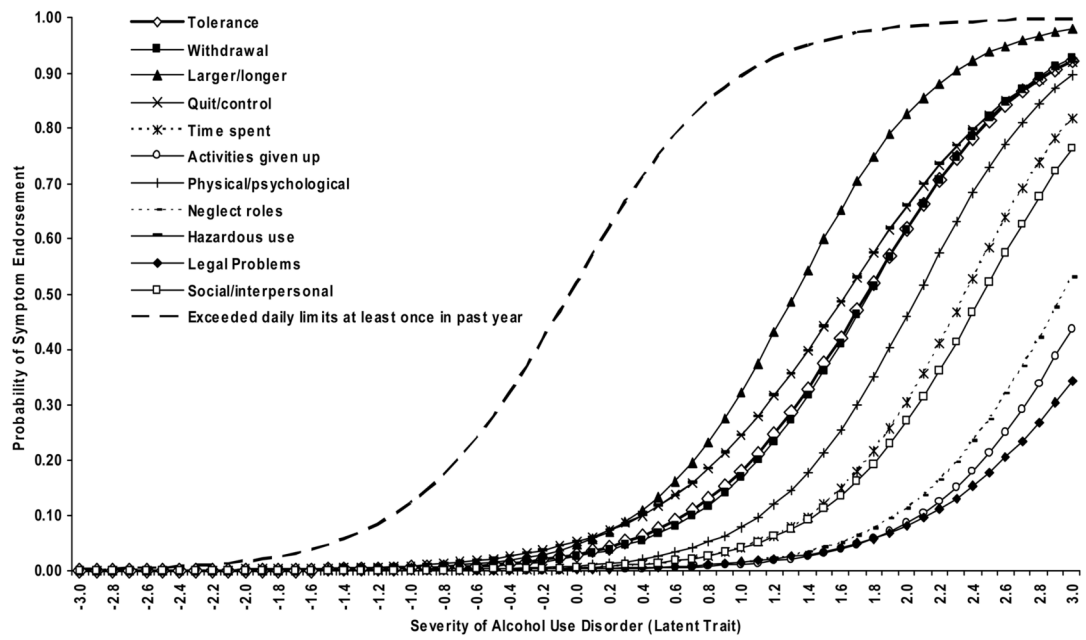
Figure 1.
Criterion response curves for DSM-IV abuse and dependence criteria and 5+/4+ drinking at least once in past year.
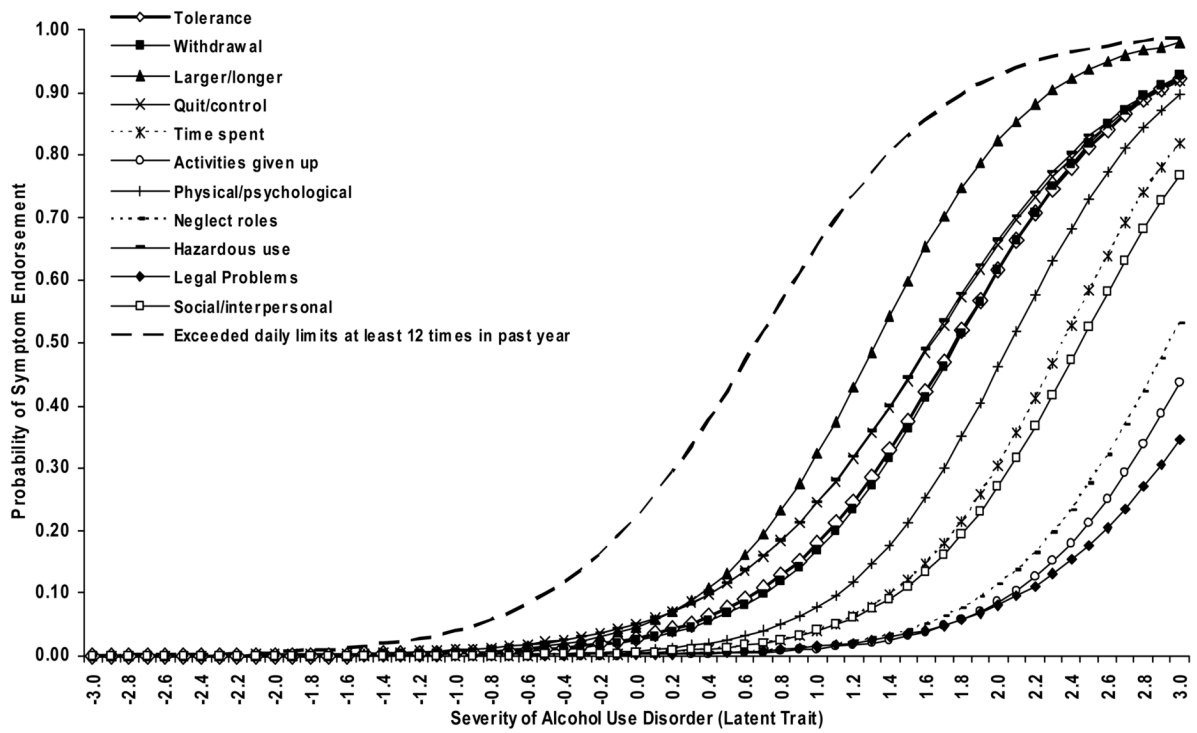
Figure 2.
Criterion response curves for DSM-IV abuse and dependence criteria and 5+/4+ drinking at least 12 times in past year.
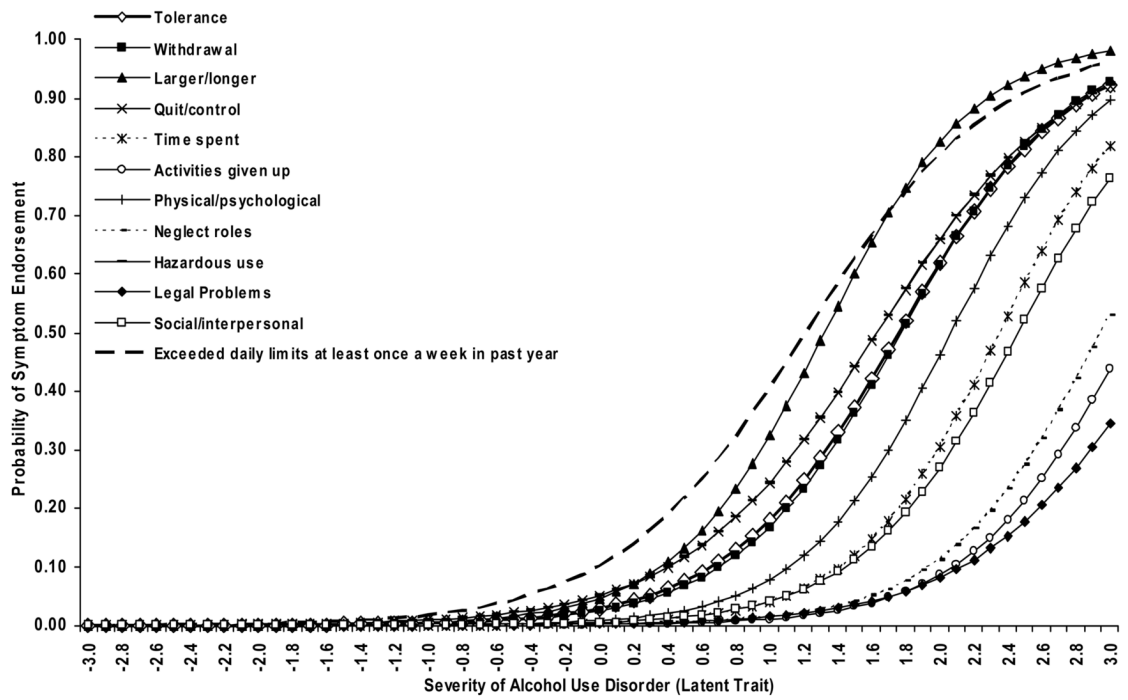
Figure 3.
Criterion response curves for DSM-IV abuse and dependence criteria and 5+/4+ driking at
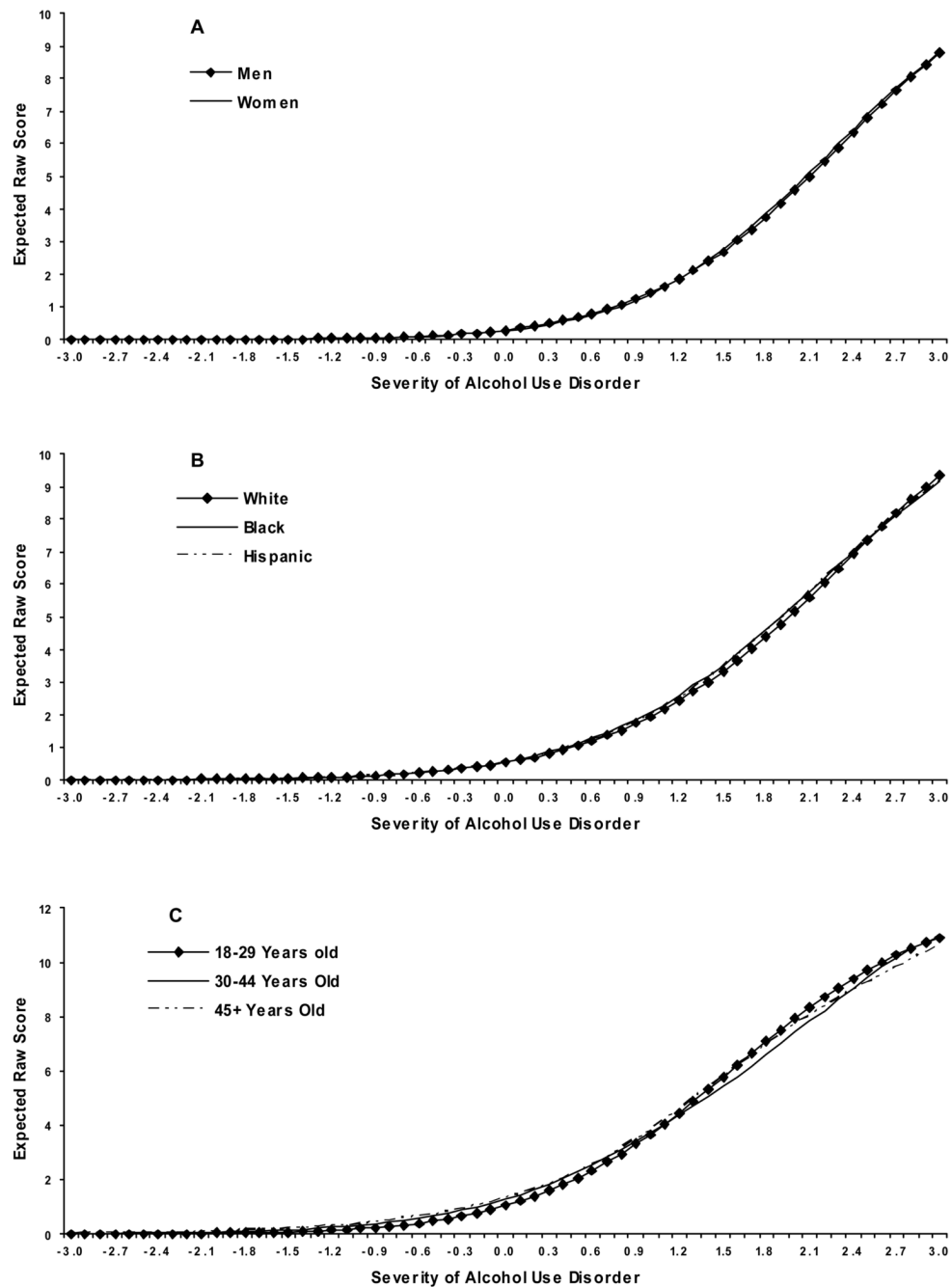least once a week in past year.

Figure 4.
Test response curves (TRCs) for sex (A), race-ethnicity (B) and age (C) for DSM-IV abuse and dependence criteria, and drinking 5+/4+ at least once a week in the past year.

Table 1

Prevalence, Factor Loadings and Criterion Response Parameters: DSM-IV Alcohol Abuse and Dependence Criteria and 5+/4+ Drinking Pattern Criteria

Past Year Alcohol Abuse (A), Dependence (D) and Consumption (C) Criteria	Prevalence (%)	Factor loadings			Model with consumption variable "Exceeded daily limits at least once in past year" Two Parameter Model		Model with consumption variable "Exceeded daily limits at least 12 times in past year" Two Parameter Model		Model with consumption variable "Exceeded daily limits at least once a week in past year" Two Parameter Model	
		With "Exceeded daily limit at least once"	With "Exceeded daily limit at least 12 times"	With "Exceeded daily limit at least once a week"	Discrimination (a) Estimate a (SE)	Severity (b) Estimate b (SE)	Discrimination (a) Estimate a (SE)	Severity (b) Estimate b (SE)	Discrimination (a) Estimate a (SE)	Severity (b) Estimate b (SE)
Tolerance (D)	9.29	0.761	0.762	0.761	1.99 (0.04)	1.76 (0.02)	1.99 (0.04)	1.76 (0.02)	1.99 (0.04)	1.76 (0.02)
Withdrawal (D)	8.80	0.803	0.800	0.799	2.10 (0.04)	1.76 (0.02)	2.07 (0.04)	1.77 (0.02)	2.06 (0.04)	1.77 (0.02)
Larger/Longer (D)	13.84	0.825	0.820	0.819	2.34 (0.05)	1.31 (0.02)	2.28 (0.05)	1.32 (0.02)	2.29 (0.05)	1.32 (0.02)
Quit/Control (D)	12.31	0.715	0.718	0.718	1.76 (0.03)	1.64 (0.02)	1.78 (0.03)	1.63 (0.02)	1.78 (0.03)	1.63 (0.02)
Time spent (D)	3.46	0.879	0.883	0.884	2.33 (0.06)	2.35 (0.03)	2.33 (0.06)	2.35 (0.03)	2.33 (0.06)	2.35 (0.03)
Activities given up (D)	1.15	0.883	0.887	0.892	2.11 (0.06)	3.11 (0.05)	2.11 (0.06)	3.12 (0.05)	2.11 (0.06)	3.12 (0.05)
Physical/Psychological problems (D)	5.26	0.883	0.884	0.884	2.32 (0.06)	2.06 (0.02)	2.31 (0.06)	2.07 (0.02)	2.30 (0.06)	2.07 (0.03)
Neglect roles (A)	1.36	0.899	0.895	0.898	2.19 (0.06)	2.94 (0.04)	2.18 (0.06)	2.95 (0.04)	2.18 (0.06)	2.95 (0.04)
Hazardous use (A)	11.56	0.718	0.715	0.712	1.82 (0.03)	1.62 (0.02)	1.81 (0.03)	1.62 (0.02)	1.79 (0.03)	1.63 (0.02)
Legal problem (A)	1.15	0.672	0.672	0.671	1.79 (0.05)	3.35 (0.06)	1.78 (0.05)	3.36 (0.06)	1.78 (0.05)	3.36 (0.06)
Social/Interpersonal problems (A)	3.09	0.865	0.868	0.865	2.18 (0.06)	2.45 (0.03)	2.19 (0.06)	2.45 (0.03)	2.17 (0.05)	2.46 (0.03)
Exceeded daily limits at least once (C)	49.15	0.730	-	-	2.06 (0.04)	−0.04 (0.01)	-	-	-	-
Exceeded daily limits at least 12 times (C)	30.18	-	0.728	-	-	-	1.91 (0.04)	0.66 (0.01)	-	-
Exceed daily limits at least once a week (C)	18.54	-	-	0.704	-	-	-	-	1.79 (0.03)	1.22 (0.02)
Bayesian Information Criterion (BIC)					105471.87		102238.10		97610.48	
Comparative Fit Index (CFI)		0.992	0.992	0.991						
Tucker Lewis Index (TLI)		0.995	0.996	0.996						

Past Year Alcohol Abuse (A), Dependence (D) and Consumption (C) Criteria	Preva-lence (%)	Factor loadings			Model with consumption variable "Exceeded daily limits at least once in past year" Two Parameter Model		Model with consumption variable "Exceeded daily limits at least 12 times in past year" Two Parameter Model		Model with consumption variable "Exceeded daily limits at least once a week in past year" Two Parameter Model	
		With "Exceeded daily limit at least once"	With "Exceeded daily limit at least 12 times"	With "Exceeded daily limit at least once a week"	Discrimination (a) Estimate a (SE)	Severity (b) Estimate b (SE)	Discrimination (a) Estimate a (SE)	Severity (b) Estimate b (SE)	Discrimination (a) Estimate a (SE)	Severity (b) Estimate b (SE)
Root mean squared error of approximation (RMSEA)		0.020	0.020	0.019						