Published in final edited form as:

Drug Alcohol Depend. 2012 July 1; 124(1-2): 88–94. doi:10.1016/j.drugalcdep.2011.12.013.

DSM-IV alcohol abuse and dependence criteria characteristics for recent onset adolescent drinkers

Jennifer S. Rose, Chien-Ti Lee, Arielle S. Selya, and Lisa C. Dierker

Department of Psychology, Wesleyan University, 207 High Street, Middletown, CT 06459

Abstract

Background—Little is known about the psychometric properties of alcohol abuse and dependence criteria among recent-onset adolescent drinkers, particularly for those who consume alcohol infrequently. This study evaluated how well DSM-IV alcohol dependence criteria measure an alcohol use disorder (AUD) construct for recent onset adolescent drinkers at different levels of drinking frequency.

Method—Data were drawn from the National Survey on Drug Use and Health, a nationally representative sample of 9,356 recent-onset adolescent drinkers, aged 12–21, who began drinking within the past year. Multiple group item response theory analysis was conducted to assess the 11 DSM-IV alcohol abuse and dependence criteria.

Results—Criteria most likely to be endorsed at lower AUD severity included ““withdrawal,” “problems at home, school or work” and “tolerance.” The criteria “drinking larger amounts/longer period of time,” “unsuccessful efforts to cut down” and “continuing to drink despite related health problems” were more likely to be endorsed at higher AUD severity. Two criteria, “tolerance” and “time spent getting, using or recovering from alcohol” showed differential item functioning between drinking frequency groups (< 7 vs. 7 days in past month), with lower discrimination and severity for more frequent drinkers. DSM-IV criteria were most precise for intermediate levels of AUD severity.

Conclusions—All but two DSM-IV criteria had consistent psychometric properties across drinking frequency groups. Symptoms were most precise for a narrow, intermediate range of AUD severity. Those assessing AUD in recent onset adolescent drinkers might consider additional symptoms to capture the full AUD continuum.

© 2011 Elsevier Ireland Ltd. All rights reserved.

Correspondence to Jennifer Rose, Wesleyan University, Department of Psychology, 207 High Street, Middletown, CT 06459. Phone: 860-685-2406. Fax: 860-685-2761. jrose01@wesleyan.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Contributors

Jennifer Rose planned the core analyses, conducted all the analyses for the revised manuscript, wrote the revised manuscript, and provided revisions to earlier drafts of this paper. Chien-Ti Lee conducted analyses for the original paper, and wrote the majority of the original paper. Arielle Selya assisted with literature searches and revisions of the earlier drafts of this paper. Lisa Dierker designed the study in conjunction with the Jennifer Rose, and provided revisions of this paper. All authors contributed to and have approved the final manuscript.

Conflict of Interest

None declared.

Keywords

alcohol dependence; drinking frequency; item response theory; recent onset drinkers; adolescents

1. Introduction

Alcohol use disorders (AUD) are a serious public health issue, often developing during adolescence (Young et al., 2002). Identifying adolescents at risk for AUD is a necessary step to reduce this problem. One of the most commonly used measures for diagnosing AUD is the DSM-IV. Traditionally, alcohol abuse and dependence are scored by the number of criteria endorsed with either a binary indicator of AUD presence or absence or a proportion score based on number of endorsed symptoms. The binary method ignores the full continuum of AUD severity that is thought to characterize alcohol dependence (Edwards and Gross, 1976). A proportion score assesses AUD severity in terms of taking into account the exact number of symptoms, yet both traditional scoring methods consider only the number of criteria endorsed, not the pattern. Consequently, it is assumed that all criteria equally contribute to the severity of AUD, and that measurement properties of the criteria are invariant across all characteristics of the individual. Finally, because DSM-IV AUD criteria were not designed specifically for adolescents, it remains unclear how the psychometric properties of these criteria vary for adolescents who have recently begun drinking and who are likely to have wide range of drinking exposure (Chung and Martin, 2005; Chung et al., 2005; Deas et al., 2000).

A number of studies have evaluated the psychometric properties of DSM-IV AUD criteria. Generally, these criteria discriminate well between levels of an underlying AUD construct, and have high probabilities of endorsement across a wide range of AUD severity. For example, among adults, the criteria “using a larger amount or over a longer period than intended” and “unsuccessful efforts to cut down” have been found to be more likely to be endorsed at low severity and the criteria “withdrawal,” “tolerance” and “activities given up or reduced” more likely to be endorsed at high AUD severity (Saha et al., 2007; Saha et al., 2006). Other IRT studies on DSM-IV criteria in adolescent drinkers have revealed low severity and/or poor discrimination for “tolerance,” “using larger amounts or over a longer period than intended” and “unsuccessful efforts to cut down” (Beseler et al., 2010; Harford et al., 2009; Hasin et al., 2003; Gelhorn et al., 2008; Martin et al., 2006). Only one study has investigated DSM-IV AUD psychometric properties in recent onset drinkers, finding that these properties differed as a function of time since onset of drinking in mostly young adults and adolescents who reported drinking in the past year and drinking onset within two years (McBride and Cheng, 2011). The psychometric properties of DSM-IV criteria have also been shown to vary as a function of age and gender (Harford et al., 2009).

Despite the fact that recent onset drinkers may have a wide range of drinking exposure, there have been no studies to date examining how the psychometric properties of DSM-IV AUD criteria differ as a function of drinking frequency. AUD symptoms are conventionally thought to emerge after prolonged and frequent exposure to alcohol (Anthony et al., 2005; Wagner and Anthony, 2002), but more recent studies have shown that symptoms can emerge within the first year of alcohol use (Lee et al., 2011; McBride and Cheng, 2011) and can predict future alcohol dependence (Behrendt et al., 2008; McBride and Cheng, 2011). Given that young recent onset drinkers are likely to drink considerably less frequently than older established drinkers on whom the DSM-IV criteria have been developed, it is important to determine whether the psychometric properties of these criteria are consistent, particularly for infrequent drinkers.

If psychometric characteristics of AUD criteria differ as a function of frequency of alcohol use, the AUD construct assessed by the same DSM-IV criteria may have a different meaning for adolescent drinkers depending on how frequently they drink. This can lead to problems in construct measurement, as well as interpretation, if potential differences in the meaning of the construct are not considered. If not considered when scoring an AUD severity construct, observed individual differences on these criteria related to drinking frequency are confounded with differences in the psychometric properties of the criteria.

The present study seeks to replicate existing studies by examining the likelihood of endorsement and discriminatory properties of DSM-IV AUD criteria, and extend this work by investigating whether criteria properties vary according to drinking frequency in a younger (ages 12–21), nationally representative population of 9,356 recent onset (within the past year) and current (drank within the past 30 days) drinkers drawn from seven consecutive years of the National Survey on Drug Use and Health (NSDUH; 2002–2008).

We used item response theory (IRT) analysis to achieve these aims. Psychometric properties of criteria can differ in: 1) severity (the value of the latent AUD construct at which the probability of item endorsement is 50%) and 2) discrimination (how well an item differentiates between participants with different levels of latent AUD). IRT analysis helps to identify possible limitations when applying DSM-IV AUD criteria to adolescents and provide guidelines for refining the DSM-IV criteria to be better suited for adolescents. IRT evaluates how well each criterion performs on a latent AUD construct and how informative each criterion is over the continuum of that construct (Gelhorn et al., 2008; Martin et al., 2006).

2. Methods

2.1 Participants

The sample consisted of $N = 9,356$ individuals ages 12–21 who reported (1) drinking in the past month and (2) their first exposure to alcohol within the past year. The NSDUH utilized multistage area probability methods to select a representative sample of the noninstitutionalized U.S. population age 12 or older. Persons living in households, military personnel living off bases, and residents of noninstitutional group quarters including college dormitories, group homes, civilians on military installations, and persons with no permanent residence are included. The NSDUH oversamples adolescents age 12–17 to improve precision of substance use estimates. In home interviews were conducted using computer-assisted interviewing and audio computer assisted self interview for sensitive questions, including substance use questions. Parental consent was required for participants ages 12–17. Participants received \$30 for participating. Weighted interview response rates ranged between 73.9% in 2007 and 79% in 2002. Data collection procedures were designed to minimize individual nonresponse bias. Nonresponse on substance use items was very low (around 1%). Weighting and imputation methods were used to adjust for nonresponse.

Half the sample was female (52.7%) with an average age of 17 years ($SE = .03$). The sample was largely non-Hispanic White (66.1%) with 15.6% Hispanic, 12.5% non-Hispanic Black, 4.0% Asian/Pacific Islander, 1.3% Interracial, and 0.5% Native Americans. The majority (87.1%) drank less than seven days in the past month. The average quantity on drinking days was 3.31 drinks ($SE = .05$) for those drinking less than 7 days in the past month and 5.26 drinks ($SE = .16$), for those drinking 7 or more days in the past month. About 13.4% met DSM-IV criteria for alcohol abuse and 8.7% met criteria for alcohol dependence.

2.2 Measures

2.2.1 Drinking frequency—Participants were asked how many days they drank in the past 30 days. Responses were dichotomized into drinking on fewer than seven days (less frequent drinkers) versus seven days or more (more frequent drinkers) to distinguish adolescent drinkers with relatively frequent drinking patterns (i.e., a total of at least one week of drinking in the past month) from the rest. Sensitivity analysis with a more liberal cut off point (< 4 vs. ≥ 4 days or more) demonstrated that changing the cutoff point did not yield different results.

2.2.2 AUD criteria—Past year alcohol abuse and dependence was assessed for participants who reported any past year use of alcohol in the 2002–2008 NSDUH with variables assessing seven DSM-IV dependence criteria (APA, 1994), including (1) tolerance, (2) withdrawal, (3) using larger amounts over a longer period than intended, (4) unsuccessful efforts to quit or cut down, (5) great deal of time spent to obtain, use or recover from drinking, (6) reduced activities, and (7) drinking despite physical or psychological problems caused by drinking; and four abuse criteria assessing 1) did something physically dangerous while under the influence of alcohol, and alcohol related problems with 2) home, school or work, 3) the law and 4) family or friends (see Table 1 for a detailed description).

2.2.3 Covariates—Because past research has identified demographic differences in criteria psychometric properties (Harford et al., 2009), age, gender, and White ethnicity covariates were included in the IRT models to remove potential confounding of mean differences on the AUD construct with differences in the psychometric properties of DSM-IV criteria. In addition, drinking quantity (average number of drinks per day in past month) was controlled to examine frequency related differences in the psychometric properties of criteria independent of drinking quantity. Quantity was recoded to 30 drinks per day for a very small number of participants reporting an excessive number of drinks (0.2%).

2.3 Analysis

Prior research supports a unidimensional AUD factor underlying the 11 DSM-IV criteria (Dawson et al., 2010; Harford et al., 2009; McBride and Cheng, 2011; T. D. Saha et al., 2006). An exploratory factor analysis confirmed this in the current sample with the majority of the common variance in the 11 DSM-IV criteria explained by a single factor. Eigenvalues for the first and second factors were 5.71 and 0.96 for less frequent drinkers (52% of variance explained by first factor) and 4.90 and 1.14 for more frequent drinkers (44.5% of variance explained by first factor). In addition, confirmatory factor analysis specifying 2 separate factors for dependence and abuse criteria did not fit the data any better than the one factor model and the two factors were highly correlated, providing further support for a single underlying AUD factor.

Single factor multiple group IRT models were tested. IRT estimates severity and discrimination for each criterion, and evaluates whether these psychometric properties are invariant across different levels of drinking frequency (adolescents who drank less than seven days in the last month versus those who drank on seven or more days; Muthén and Lehman, 1985), while controlling for other covariates (i.e., age, gender, white ethnicity, drinking quantity).

Severity is a function of the criterion's threshold. A criterion with a low threshold has low severity, indicating that it is more likely to be endorsed at lower levels of the AUD construct. Discrimination is a function of the slope (factor loading) of the criterion; the steeper the slope, the better the criterion discriminates. Item characteristic curves (ICC) were

plotted to examine criteria severity and discrimination simultaneously at different values of the latent AUD construct.

Differential item functioning (DIF), that is whether DSM-IV criteria have invariant psychometric properties across levels of drinking frequency, was also investigated. Invariance suggests that the latent AUD construct is equivalent at different levels of frequency (Teresi and Fleishman, 2007). Criteria that differ in severity and/or discrimination across drinking groups exhibit DIF. A criterion with severity DIF indicates that adolescents with different levels of drinking frequency have a different probability of endorsing that criterion despite having the same level of latent AUD, and criteria with discrimination DIF have different levels of discrimination across drinking frequency groups.

Ten multiple group IRT models were tested to identify criteria with DIF while controlling for mean differences on the latent AUD construct as a function of age, gender, white ethnicity, and drinking quantity. Uncontrolled mean differences in the latent AUD construct can appear as item DIF in the IRT models. Thus, we controlled for these characteristics to better isolate drinking frequency related DIF from factor mean differences. Each model compared a constrained model (where slope and threshold parameters of a chosen criterion were constrained to be equal across drinking frequency groups) and an unconstrained model (in which slope and threshold parameters of all criteria were free to vary across groups, with the exception of a single anchor criterion). The anchor criterion was chosen from preliminary factor analysis based on: (1) high factor loading on the AUD construct in both groups, and (2) no evidence of DIF (Stark et al., 2006). An approximate chi-square difference test for each comparison provides a *p* value for the difference in model fit (adjusted for multiple significance tests; Benjamini and Hochberg, 1995) where a significant adjusted *p* value indicates DIF.

Finally, we tested a final IRT model in which all criteria showing DIF were free to vary across groups, from which we generated ICCs and calculated a model estimated maximum a posteriori (MAP) AUD severity factor score recommended by Thissen and Wainer (2001). The MAP score, calculated for each observation, is based on the mode of posterior distribution of the model parameter estimates, and takes into account each individual's response pattern and differences in IRT parameter estimates. All analyses used sample weights to correct for selection probabilities, and adjusted for survey design effects to obtain accurate standard errors.

3. Results

IRT results showed that severity and discrimination parameters for all but two criteria were invariant across drinking frequency groups after controlling for age, gender, ethnicity, and drinking quantity. The criteria "tolerance" and "time spent" had lower severity and did not discriminate as well for more frequent drinkers. The final multiple group IRT model was one in which all criteria slopes and thresholds were constrained to be equal between the two groups of drinkers, except for the two noninvariant criteria. Fit indices indicated good fit, TLI = .95, CFI = .96, RMSEA = .016. Additionally, the *p*-value associated with the approximate change in chi-square for the final multiple group IRT model compared with the unconstrained model was not significant (*p* = .26), indicating that the model fit was not significantly worse by constraining 9 of the 11 criteria across groups.

For both drinking frequency groups, being female (*B* = .12 and .29 for less and more frequent drinkers, respectively) and younger (*B* = -.04 and -.10 for less and more frequent drinkers, respectively) were significantly associated with higher AUD severity factor scores (all *p* < .05). Greater drinking quantity was significantly associated with higher AUD

severity scores for less frequent drinkers ($B = .07, p = .000$), but not for more frequent drinkers ($B = .02, p = .06$). White ethnicity was not significantly associated with AUD severity in either group.

Table 1 shows severity and discrimination estimates from the final model, and ICCs for drinking frequency groups are presented in Figures 1 and 2. Both “tolerance” and “time spent” had poorer discrimination among more frequent drinkers compared to less frequent drinkers. Criteria that best discriminated among levels of AUD for both groups of adolescent drinkers were “reduce activities” and “family/friend problems,” “home problems” and “health problems.” The criteria “cut down,” “withdrawal” and “larger/longer” showed the poorest discrimination. Criteria with low severity included “withdrawal,” “home/school/work problems” and “tolerance” (particularly for less frequent drinkers), suggesting that these criteria were more likely to be endorsed at lower levels of the AUD construct. Conversely, criteria more likely to be endorsed at higher levels of AUD construct included “larger/longer,” “cut down” and “health problems.”

Figure 3 presents the total information curve for all 11 DSM-IV criteria for both groups of adolescent drinkers; it shows the amount of information contained by the criteria across the underlying AUD continuum. The peak information value indicates more precise measurement of the latent AUD construct at the corresponding factor score (Muthén, 1996). Accordingly, the 11 criteria for both groups were most precise at measuring intermediate levels of AUD severity. A higher peak for less frequent drinkers suggested that the criteria as a whole were somewhat less precise for more frequent adolescent drinkers.

The distribution of AUD MAP scores for less frequent drinkers (ranging from -1.18 to 2.02) was positively skewed, with more than half having a score below the mean (Figure 5a). For more frequent drinkers, the distribution was closer to a normal distribution (ranging from $-.91$ to 2.10) with more adolescents having higher levels of AUD severity based on MAP scores (Figure 5b). Figure 6a and 6b show scatterplots of the association between an AUD score based on proportion of endorsed symptoms and MAP scores for both drinking frequency groups. Although both scores were highly correlated ($r = .94$ and $.97$ for less and more frequent drinkers), the MAP scores had significantly more variance as evidenced by the dispersion of MAP scores at each level of the proportion score. This is because the MAP score takes into consideration not only the number of criteria endorsed, but also which criteria were endorsed, and this allows for a more precise estimate of AUD severity.

4. Discussion

This study evaluated how DSM-IV alcohol abuse and dependence criteria measured an underlying AUD construct and whether individual criteria function differently based on the level of past month drinking frequency among recent-onset adolescent drinkers. A few major findings emerged. First, the majority of the criteria, with the exception of “tolerance” and “time spent,” showed good discrimination and were invariant in terms of both discrimination and severity across levels of drinking frequency, after controlling for age, gender, white ethnicity, and drinking quantity. This suggests that these criteria were strongly associated with the underlying AUD construct, regardless of the frequency of alcohol consumption.

Second, criteria more likely to be endorsed at low severity included withdrawal, home/school/work problems and tolerance. Tolerance and withdrawal are symptoms of alcohol dependence traditionally thought to be associated with heavier and more prolonged drinking. There is evidence to suggest that adolescents’ progression to alcohol dependence is significantly more rapid than adults (Martin et al., 1995). These results support this finding,

and suggest that these symptoms may be among those experienced by adolescents long before prolonged and heavy drinking occurs. In fact, tolerance was the most endorsed criterion (20.2%) among less frequent drinkers. Although the finding that tolerance tended to be endorsed at lower levels of the AUD construct is inconsistent with findings from adults and treatment seeking adolescents, it is consistent with new research on the psychometric properties of DSM-IV criteria in drinkers who began drinking in the past two years (McBride and Cheng, 2011). Taken together, this research suggests that this finding may be more specific to those who have just begun drinking and drink considerably less frequently than more established drinkers. Rather than representing a physiological tolerance for alcohol, the high endorsement rate for tolerance among more frequent, yet still relatively light, adolescent drinkers may be more representative of the natural development of adolescent drinking, which typically starts with small quantities consumed infrequently, and increases over time (Chung and Martin, 2005).

Symptoms with high AUD severity included “larger/longer,” “health problems” and “cut down.” The low endorsement rates for “larger/longer” and “health problems” are consistent with rates found for young, recent onset drinkers from nationally representative and community samples (Gelhorn et al., 2008; Harford et al., 2009; McBride and Cheng, 2011). On the other hand, the finding that the criterion “cut down” had high AUD severity and was one of the least endorsed symptoms (2.3% and 6.9% of less and more frequent drinkers, respectively, endorsed this symptom) is inconsistent with previous findings that this criterion was the most frequently endorsed criterion with 12.6% endorsing this symptom with 90 days of beginning drinking (McBride and Cheng, 2011). One reason may be a difference in defining the criterion. We used a conservative definition of having been unsuccessful every time an attempt was made to quit or cut down in the past 12 months. Other ways of defining this criterion including questions about wanting to quit or cut down, setting limits or having attempted to quit or cut down may produce higher prevalence rates.

IRT severity estimates are associated with criteria base rates. Criteria with low endorsement rates are often endorsed by those who have also endorsed a greater number of criteria overall. Because an individual's score on the underlying AUD construct is based on both number of criteria endorsed and psychometric properties of the criteria, a greater number of endorsed criteria results in greater AUD severity. Therefore, criteria like “cut down” that are endorsed by fewer drinkers tend to have a higher probability of being endorsed at greater AUD severity.

Unlike previous studies that found DSM-IV AUD criteria to be most precise at high levels of the AUD construct for adolescents undergoing treatment (Gelhorn et al., 2008; Martin et al., 2006), our findings showed that these criteria were most precise for measuring moderate AUD. This may be because fewer adolescents in our study endorsed criteria with high severity, given that they were recent-onset drinkers with a shorter history of alcohol use and a lower amount of alcohol consumption compared to adolescents from treatment settings. Additionally, the information curve peak was higher for less frequent drinkers, implying that DSM-IV criteria were more precise for this group than for more frequent drinkers. This is likely due to high endorsement of tolerance and time spent among more frequent drinkers (Figure 4). Although they were the most frequently endorsed criteria for both less frequent (20.0% and 11.8%, for tolerance and time spent) and more frequent adolescent drinkers (50.5% and 45.1% for tolerance and time spent), endorsement rates were much higher among more frequent drinkers.

Overall, the DSM-IV alcohol abuse and dependence criteria appeared to have good, stable psychometric properties across levels of drinking frequency, but there were differences in the psychometric properties of tolerance and time spent between drinking frequency groups.

Although lack of invariance at different levels of drinking frequency for tolerance and time spent may make them seem less desirable from a psychometric standpoint, they were the most frequently endorsed symptoms in both drinking frequency groups. Therefore, they may serve as important indicators of developing AUD in adolescents regardless of how infrequently they are drinking at the time of measurement. However, lack of invariance suggests that those assessing DSM-IV criteria in adolescent drinkers should be aware of the potential for confounding of differences in psychometric properties with true differences in AUD, and might consider using a factor score that accounts for DIF.

Other low severity criteria that might serve as early indicators of AUD development include withdrawal and home/school/work problems, despite their lower endorsement rates among adolescent drinkers. On the other hand, criteria reflecting unsuccessful attempts to manage drinking as evidenced by drinking more than intended and an inability to cut down or quit drinking could serve as indicators of more advanced AUD in adolescents. Finally, other potential AUD symptoms tapping lower levels of AUD might be used in addition to DSM-IV criteria for recent-onset drinkers to ensure that lower AUD severity levels are captured. For example, expressing a desire to cut down on drinking may reflect a readiness to change that could signal onset of low severity AUD and a potential opportunity for early intervention. However, it remains to be seen in future research whether the addition of this kind of symptom, or the increased precision gained by using an AUD factor score, can more accurately predict important outcomes such as chronic drinking, cessation and relapse in adolescent drinkers.

A major strength of this study is that the large nationally representative dataset provided large enough sample sizes and statistical power to evaluate the DSM-IV criteria across levels of drinking frequency, and to generalize more confidently to a broader population of recent-onset adolescent drinkers. Additionally, our study was the first to identify drinking frequency DIF in AUD criteria in an understudied population of recent onset adolescent drinkers, while controlling for age, gender, ethnicity, and drinking quantity. Some limitations should also be noted. First, we could not identify reasons for observed differences in the psychometric properties of the criteria. It is unclear to what extent adolescent self reported endorsement of AUD criteria may be complicated by interpretations that differ from adults. Second, neither IRT nor the cross-sectional data can inform us about predictive validity. Third, IRT does not inform us about the mechanisms underlying variability in patterns of criteria endorsement. Despite these limitations, this study provides support for psychometric stability in most DSM-IV criteria across levels of drinking frequency in novice adolescent drinkers, and suggests that future measurement of alcohol dependence in this population might benefit from including other symptoms that better capture lower AUD severity.

Acknowledgments

Role of funding resources

Data analysis and interpretation, manuscript preparation, review and approval were supported by grants R01 DA022313-01A2, R01 DA022313-02S1, R21DA024260, (Dierker), R21DA029834-01A1 (Rose) from the National Institute on Drug Abuse, and P50 DA010075 awarded to Penn State University and an Investigator Award from the McManus Charitable Trust (Dierker). We acknowledge The National Household Survey of Drug Use and Health (NHSDUH) funded by The Substance Abuse and Mental Health Services Administration (SAMHSA), U.S. Public Health Service in the [U.S. Department of Health and Human Services](#) (DHHS).

References

American Psychology Association. Diagnostic and Statistical Manual of the Mental Disorders - IV. Washington, D.C.: APA; 1994.

Drug Alcohol Depend. Author manuscript; available in PMC 2013 July 01.

- Anthony JC, Chen CY, Storr CL. Drug dependence epidemiology. *Clin. Neurosci. Res.* 2005; 55:55–68.
- Behrendt S, Wittchen HU, Hofler M, Lieb R, Low NC, Rehm J, Beesdo K. Risk and speed of transitions to first alcohol dependence symptoms in adolescents: a 10-year longitudinal community study in Germany. *Addiction.* 2008; 103:1638–1647. [PubMed: 18821874]
- Benjamini Y, Hochberg Y. Controlling the false discovery rates: a practical and powerful approach to multiple significance testing. *J. Royal Stat. Soc.* 1995; 57:289–300.
- Beseler CL, Taylor LA, Leeman RF. An item-response theory analysis of DSM-IV alcohol-use disorder criteria and "binge" drinking in undergraduates. *J. Stud. Alcohol Drugs.* 2010; 71:418–423. [PubMed: 20409436]
- Chung T, Martin CS. What were they thinking? Adolescents' interpretations of DSM-IV alcohol dependence symptom queries and implications for diagnostic validity. *Drug Alcohol Depend.* 2005; 80:191–200. [PubMed: 15894432]
- Chung T, Martin CS, Winters KC. Diagnosis, course, and assessment of alcohol abuse and dependence in adolescents. *Rec. Dev. Alcohol.* 2005; 17:5–27.
- Dawson DA, Saha TD, Grant BF. A multidimensional assessment of the validity and utility of alcohol use disorder severity as determined by item response theory models. *Drug Alcohol Depend.* 2010; 107:31–38. [PubMed: 19782481]
- Deas D, Riggs P, Langenbucher J, Goldman M, Brown S. Adolescents are not adults: developmental considerations in alcohol users. *Alcohol. Clin. Exp. Res.* 2000; 24:232–237. [PubMed: 10698377]
- Edwards G, Gross MM. Alcohol dependence: provisional description of a clinical syndrome. *BMJ.* 1976; 1:1058–1061. [PubMed: 773501]
- Gelhorn H, Hartman C, Sakai J, Stallings M, Young S, Rhee SH, Corley R, Hewitt J, Hopfer C, Crowley T. 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.* 2008; 47:1329–1339. [PubMed: 18827724]
- Harford TC, Yi HY, Faden VB, Chen CM. 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.* 2009; 33:868–878. [PubMed: 19320629]
- 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]
- Lee C-T, Rose JS, Engel-Rebitzer E, Selya A, Dierker L. Alcohol dependence symptoms among recent onset adolescent drinkers. *Addict. Behav.* 2011; 36:1160–1167. [PubMed: 21835550]
- Martin CS, Chung T, Kirisci L, Langenbucher JW. Item response theory analysis of diagnostic criteria for alcohol and cannabis use disorders in adolescents: implications for DSM-V. *J. Abnorm. Psychol.* 2006; 115:807–814. [PubMed: 17100538]
- Martin CS, Kaczenski NA, Maisto SA, Bukstein AM, Moss HB. Patterns of DSM-IV alcohol abuse and dependence symptoms in adolescent drinkers. *J. Stud. Alcohol.* 1995; 56:672–680. [PubMed: 8558899]
- McBride O, Cheng HG. Exploring the emergence of alcohol use disorder symptoms in the two years after onset of drinking: findings from the National Surveys on Drug Use and Health. *Addiction.* 2011; 106:555–563. [PubMed: 21182549]
- Muthén BO. Psychometric evaluation of diagnostic criteria: application to a two-dimensional model of alcohol abuse and dependence. *Drug Alcohol Depend.* 1996; 41:101–121. [PubMed: 8809498]
- Muthén BO, Lehman J. Multiple group IRT modeling: applications to item bias analysis. *J. Educ. Stat.* 1985; 10:133–142.
- Saha T, Stinson FS, Grant BF. The role of alcohol consumption in future classifications of alcohol use disorders. *Drug Alcohol Depend.* 2007; 89:82–92. [PubMed: 17240085]
- 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]

- Stark S, Chernyshenko OS, Drasgow F. Detecting differential item functioning with confirmatory factor analysis and item response theory: toward a unified strategy. *J. Appl. Psychol.* 2006; 91:1292–1306. [PubMed: 17100485]
- Teresi JA, Fleishman JA. Differential Item functioning and health assessment. *Qual. Life Res.* 2007; 16:33–42. [PubMed: 17443420]
- Thissen, D.; Wainer, H. *Test Scoring*. Mahwah, NJ: Erlbaum; 2001.
- Wagner FA, Anthony JC. From first drug use to drug dependence: developmental periods of risk for dependence upon marijuana, cocaine, and alcohol. *Neuropsychopharmacology*. 2002; 26:479–488. [PubMed: 11927172]
- Young SE, Corley RP, Stallings MC, Rhee SH, Crowley TJ, Hewitt JK. Substance use, abuse and dependence in adolescence: prevalence, symptom profiles and correlates. *Drug Alcohol Depend.* 2002; 68:309–322. [PubMed: 12393225]

Figure 1. Item characteristic curves for DSM-IV AUD criteria for adolescents who drank less than 7 days in the past 30 days (DIF symptoms have dashed curves).

Figure 2.

Item characteristic curves for DSM-IV alcohol dependence criteria for adolescents who drank 7 days or more in the past 30 days (DIF symptoms have dashed curves).

Figure 3. Total information curve of *DSM-IV* criteria stratified by drinking frequency in the past 30 days.

Figure 4. DSM-IV Tolerance and time spent criteria information curves stratified by drinking frequency in the past 30 days.

5a

5b

Figure 5.

- a. AUD factor scores for adolescents who drank less than 7 days in the past 30 days.
- b. AUD factor scores for adolescents who drank 7 or more days in the past 30 days.

6a

6b

Figure 6.

- Scatterplot of proportion score with MAP score for adolescents who drank less than 7 days in the past 30 days.
- Scatterplot of proportion score with MAP score for adolescents who drank 7 or more days in the past 30 days

Table 1
Design adjusted symptom endorsement rates and IRT parameter estimates from the final model.

DSM-IV Criterion	Number (%) endorsing each symptom		Discrimination	Severity	$\Delta\chi^2$ (p value)
	< 7 days	$\geq$ 7days			
During the past 12 months:					
“needed to drink more than you used to in order to get the effect you wanted” and/or “drinking the same amount had less effect that it used to” (TOLERANCE)	1654 (20.0)	595 (50.5)	.92 (.04) [.61 (.06)] ^b	1.03 (.16) [.73 (.42)] ^b	14.10 (.001) [*]
“did you have 2 or more symptoms at the same time that lasted for longer than a day after you cut back or stopped drinking” (WITHDRAWAL)	415 (5.0)	151 (13.0)	.77 (.04)	.59 (.23)	2.24 (.33)
“unable to keep to limits set on drinking or often drank more than you intended to” (LARGER/LONGER)	193 (2.2)	97 (6.8)	.76 (.04)	3.31 (.31)	3.24 (.20)
“unable to cut down or stop drinking every time you wanted to or tried to” (CUT DOWN)	185 (2.0)	81 (5.5)	.68 (.05)	3.33 (.27)	4.24 (.12)
“a month or more when you spent a lot of your time getting or drinking alcohol” and/or “a month or more when you spent a lot of time getting over the effects of alcohol” (TIME SPENT)	984 (11.8)	542 (45.1)	1.13 (.04) [.85 (.06)] ^b	1.13 (.25) [−.06 (.45)] ^b	11.15 (.004) [*]
“drinking alcohol caused you to give up or spend less time doing important activities” (REDUCE ACTIVITIES)	256 (3.2)	152 (11.4)	1.31 (NA) ^a	1.52 (.28)	NA ^a
“continued to drink even though it causing you to have problems with your emotions, nerves, or mental health” and/or “continued to drink even though you thought it was causing you to have physical problems (HEALTH PROBLEMS)	199 (2.3)	134 (10.5)	1.14 (.03)	2.58 (.29)	1.66 (.44)
“drinking caused you to have serious problems at home, work or school” (HOME/SCHOOL/WORK PROBLEMS)	253 (3.4)	139 (10.7)	1.14 (.04)	.75 (.27)	1.13 (.57)
“regularly drank and then did something where being drunk might have put you in physical danger” (PHYSICAL DANGER)	535 (6.1)	288 (23.4)	.95 (.03)	1.40 (.24)	8.38 (.02)
“drinking caused you to do things that repeatedly got you in trouble with the law” (LEGAL PROBLEMS)	106 (1.1)	75 (5.5)	.85 (.05)	1.55 (.34)	7.86 (.02)
“continued to drink even though you thought it caused problems with family or friends” (FAMILY/FRIEND PROBLEMS)	227 (2.6)	132 (8.4)	1.21 (.03)	1.11 (.31)	2.90 (.23)

^aNA = anchor item initially constrained to be equal across groups for model identification

^bparameter estimate for adolescents who drank weekly or more (≥7 days)

^{*}significant using the Benjamini-Hochberg procedure for Type I error rate adjustment for multiple tests (Benjamini and Hochberg, 1995)