An Item-Response Theory Analysis of DSM-IV Alcohol-Use Disorder Criteria and "Binge" Drinking in Undergraduates*

CHERYL L. BESELER, PH.D.,† LAURA A. TAYLOR, B.A.,† AND ROBERT F. LEEMAN, PH.D.†

Department of Psychiatry, New York State Psychiatric Institute, Columbia University, 1051 Riverside Drive Box 123, New York, NY 10032

ABSTRACT. Objective: This is the first study to examine the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV), criteria for alcohol-use disorders and heavy episodic (or "binge") drinking in a college sample using item-response theory (IRT) analysis. IRT facilitates assessment of the severity of the criteria, their ability to distinguish between those at greatest and lowest risk, and the value of adding a "binge" drinking criterion. **Method:** In a sample of undergraduate drinkers ($n = 353$), we conducted factor analyses to determine whether the criteria best fit a one- or two-factor structure. We then conducted IRT analyses to obtain item-characteristic curves indicating the probability of endorsing a criterion at increasing levels of alcohol-use-disorder risk. These analyses were first conducted including current (i.e., past-year) DSM-IV alcohol-use-disorder criteria only and then rerun adding weekly "binge" drinking. **Results:** A single-factor model of the DSM-IV criteria did not differ significantly from a two-factor model. After including "binge" drinking, two factors fit the data slightly better than one factor but with a dominant single factor. Withdrawal was the most severe criterion, whereas tolerance and "larger/longer" were the least severe. Time spent drinking and a combined social/legal difficulties criterion had the best ability to discriminate those at greatest and lowest risk for an alcohol-use disorder. "Binge" drinking showed both low discrimination and low severity. **Conclusions:** To our knowledge, this is the first study to examine DSM-IV criteria in an undergraduate sample using IRT. In this sample, abuse and dependence were intermixed on a continuum of severity, and "binge" drinking was in the least severe region. (*J. Stud. Alcohol Drugs, 71,* 418-423, 2010)

GIVEN THE COMMONALITY OF ALCOHOL USE in young adults (Johnston et al., 2003; O'Malley and Johnston, 2002), it is not surprising that diagnoses of alcohol-use disorders (AUDs; i.e., alcohol abuse or dependence) are relatively prevalent in this age group. In the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC), college students ages 18-29 were almost twice as likely as adults 30 years of age or older to meet criteria for current alcohol abuse and more than four times as likely to meet criteria for current dependence (Dawson et al., 2004, Grant et al., 2004). Compared with their non-college-attending peers, college students evidence heavy episodic drinking more often and are more likely to be diagnosed with alcohol abuse (Slutske et al., 2004; Slutske, 2005).

Alcohol use among undergraduates is often typified by heavy episodic use, a pattern that increases risk for negative consequences such as academic failure and unsafe sexual practices (Jennison, 2004; Perkins, 2002; Wechsler and Nelson, 2001). In a national survey, just under 44% of undergraduates in the United States reported at least one heavy episodic drinking occasion (i.e., five or more drinks for men, four or more for women) in the prior 2 weeks (Wechsler et al., 2002). Heavy episodic drinking is closely associated with AUD diagnoses. In a national survey, frequent heavy episodic drinkers (i.e., three or more heavy drinking occasions in the prior 2 weeks) were 13 times more likely to meet criteria for current abuse and 19 times more likely to meet criteria for current dependence than drinkers not reporting heavy use (Knight et al., 2002).

Despite evidence that heavy drinking increases the likelihood of an AUD, alcohol consumption has never been a criterion in the Diagnostic and Statistical Manual of Mental Disorders (DSM) classifications of AUDs, but this could change in DSM-V (Li et al., 2007a; Li et al., 2007b; Saha et al., 2007). Analyses of the severity of the DSM-IV (American Psychiatric Association, 2000) AUD criteria and their ability to distinguish between higher and lower risk drinkers have noted that the majority of the criteria are relevant primarily to more severely dependent drinkers (Proudfoot et al., 2006; Saha et al., 2006; Saha et al., 2007). Evidence suggests that a heavy drinking criterion would tap the lower end of the severity spectrum (Saha et al., 2007) and would help to identify those who drink at heavy, potentially problematic levels, but who do not meet the criterion for tolerance (Martin et al., 2008).

Received: July 17, 2009. Revision: September 20, 2009.

*This research was supported by National Institute on Alcohol Abuse and Alcoholism grant R01 AA016621.

†Correspondence may be sent to Cheryl L. Beseler at the above address or via email at: beselerc@nypdrat.cpmc.columbia.edu. Laura A. Taylor is with Scripps College, Claremont, CA. Robert F. Leeman is with Yale University School of Medicine, New Haven, CT.

Item response theory (IRT) is used to describe the probability of observing a particular pattern of item endorsement, given an individual's underlying level of vulnerability. In a 2-parameter logistic (2PL) model for binary items, the probability is modeled using a slope parameter (discrimination) and a location parameter (difficulty) for each item (Gillespie et al., 2007). The slope parameter measures the strength of the relationship between the item and the underlying construct; a steeper slope means better item discrimination than a flatter slope on the underlying severity spectrum. The location parameter is the point on the continuum where the respondent has a 50% chance of endorsing the item.

Studies in adult samples have shown that DSM-IV alcohol dependence and abuse criteria constitute a single factor that accounts for a majority of the variance in the criteria (Kahler and Strong, 2006; Krueger et al., 2004; Proudfoot et al., 2006; Saha et al., 2006) and forms a linear continuum of severity using Rasch (Kahler et al., 2006), IRT (Krueger et al., 2004), and a method that tests for discontinuities in the criteria (Hasin and Beseler, 2009). However, there is evidence that not all the DSM-IV criteria form a severity continuum (see Langenbucher et al., 2004; Saha et al., 2006). Saha et al. (2007) included a heavy episodic drinking criterion with the AUD criteria in an IRT analysis and found that such drinking represented a low level of severity and that the linear continuum of severity observed with the AUD criteria only was maintained after this item was added. In contrast, Hasin and Beseler (2009) found that adding a heavy episodic drinking criterion led to a deviation from the linear continuum of severity formed by the AUD criteria.

There have been fewer studies on the dimensionality of DSM-IV abuse and dependence criteria in undergraduates, despite the overall breadth and depth of the literature on college drinking. In a Rasch model involving 18- to 24-year-olds in the NESARC, tolerance was less severe in older adults than in younger adults (Kahler and Strong, 2006). The present study is an examination of the unidimensionality of DSM-IV criteria for current abuse, dependence, and "binge" drinking (i.e., endorsement of drinking five or more drinks for men and four or more for women within a 2-hour period), using survey data from undergraduates. To our knowledge, this study is the first to use IRT to assess the DSM-IV AUD criteria in a sample of undergraduates.

Method

Sample and procedures

The survey was conducted at a moderate-sized state university in the northeast United States, beginning in the fall of 2007. More than 70% of the student body is White, with 28% of students housed on campus. The survey was advertised via flyers, via announcements during psychology classes, and through a Web site maintained by the Psychology Department. Participants completed the anonymous, Web-based survey in exchange for credit toward completion of a research-participation requirement for introductory psychology courses. Participants were at least 18 years of age and volunteered to participate. A history of alcohol consumption was not required.

Measures

Endorsement of abuse and dependence criteria within the past year was assessed using an 11-item self-report measure based on the DSM-IV criteria. There were two main differences between the DSM-IV criteria and the items in this measure. Two items concerning withdrawal were posed: one concerning experience of withdrawal and the other concerning alcohol use to avoid withdrawal. These two items were combined for analysis, such that endorsement of either indicated withdrawal. Also, the abuse criteria concerning legal and social/interpersonal problems were combined into one item. This combination was done to address a concern that some students may consider university-implemented penalties for drinking to be legal problems, whereas others may consider them social/interpersonal consequences—potentially threatening the validity of each item if kept separate. A copy of the measure and results of analyses to establish its validity may be obtained from the authors.

Participants reported the number of times in the past 3 months that they had consumed five or more alcoholic beverages (for men) or four or more (for women) in a 2-hour period. This is the definition of "binge drinking" put forth by the National Institute on Alcohol Abuse and Alcoholism (2004). To create a dichotomous version for analysis, participants reporting such drinking at least once per week were considered to be in this category.

The survey also included demographic items, including gender, age, race/ethnicity, and university class standing. Race/ethnicity and class standing were omitted during one semester of data collection, and thus this information was not available for a portion of the sample.

Analyses

Chi-square analyses were used to compare endorsement of each DSM-IV criterion and the "binge"-drinking criterion by gender. We used marginal maximum likelihood for the factor analysis in Mx, which maximizes the likelihood of the data conditioned on the latent trait and estimates the tetrachoric correlation matrix (Neale et al., 2003). We first factor analyzed the 10 dependence and abuse criteria and tested the fit of one- and two-factor models using the maximum likelihood ratio test, the Akaike's Information Criterion, and the Bayesian Information Criteria (smaller values indicate a better fit). We next added weekly "binge" drinking and retested the model fit. We used the item factor loadings, re-

siduals, and thresholds to calculate a two-parameter normal ogive item-response model and item-characteristic curves to display the results (Gillespie et al., 2007).

Results

Participants

A total of 425 participants completed at least half of the survey and comprised the valid sample for the larger study. Of these, 353 participants responded to all of the items in the self-report measure concerning the AUD criteria, had "binge" drinking information, and reported ever having had at least one alcoholic drink. These participants made up the sample for the present study. The sample was 73.1% female and 26.6% male, with one participant not reporting gender. Significant gender differences existed in the endorsement of only one of the DSM-IV criteria and the "binge"-drinking criterion. Men (19.1%) were more likely than women (10.5%) to endorse hazardous use, $\chi^2(1, n = 352) = 4.66$, $p = .031$, and men (54.3%) were more likely than women (41.1%) to endorse "binge" drinking, $\chi^2(1, n = 352) = 4.84$, $p = .028$. Given the low number of men ($n = 94$), differential item functioning was not conducted. The mean age was 19.03 ($SD = 2.13$, range: 18-47). Race/ethnicity and class standing were available for about half of the sample ($n = 176$ and 179, respectively). This portion of the sample was 85.2% White, 6.8% African American, 4.5% Hispanic, 1.1% Asian, and 2.3% "other." Freshman was the most common class standing (78.8%), followed by sophomore (16.8%), junior (2.8%), senior (1.1%), and "other" (0.6%). Alcohol-dependence criteria were met by 19.3% of the sample.

Prevalence of criteria and factor analysis results

Tolerance and heavy episodic ("binge") drinking were the criteria endorsed most often; the combined social/legal problems item, reducing activities to drink, and withdrawal were the least frequently endorsed criteria (Table 1). A single factor adequately fit the 10 dependence and abuse criteria, $\chi^2(9) = 16.8$, $p = .052$; eigenvalues 6.22 and 0.89 (results not shown, available on request). After adding "binge," a two-factor model was significantly preferred over a one-factor model, $\chi^2(10) = 23.0$, $p = .01$; eigenvalues 6.47 and 1.14. Akaike's Information Criterion and Bayesian Information Criteria statistics indicated that one- and two-factor models fit the 11 criteria equally well (Table 1). Adding "binge" drinking had little effect on the magnitude of the DSM-IV item factor loadings (data not shown). Given the fit statistics, we concluded that we had a single dominant latent factor of the 11 items that could be modeled using IRT.

Results of item-response theory analysis of dependence, abuse, and "binge" drinking criteria

Comparing DSM-IV and "binge" drinking items indicated that those who endorsed withdrawal showed the greatest AUD severity and those who endorsed tolerance showed

Table 1. Factor loadings with 95% confidence intervals, threshold, discrimination and difficulty parameters from one-factor Mx marginal maximum likelihood analysis of 10 alcohol dependence, abuse, and weekly "binge" drinking criteria in 353 undergraduate students, 2007

Alcohol criteria	Prevalence *n* (%)	Factor loading One-factor model [95% CI]	Factor loading Two-factor model		Discrim.	Severity
			Factor 1	Factor 2		
Tolerance	161 (45.6)	.83 [.73, .90]	.759	.284	1.461	0.127
Withdrawal	25 (7.1)	.55 [.36, .71]	.460	.350	0.661	2.667
Larger/longer	123 (34.8)	.76 [.66, .84]	.813	.041	1.164	0.509
Quit/cut down	34 (9.6)	.76 [.63, .86]	.738	.250	1.182	1.701
Time spent drinking	51 (14.4)	.86 [.77, .92]	.892	.166	1.694	1.222
Activities reduced	22 (6.2)	.78 [.64, .88]	.666	.468	1.249	1.963
Use despite problems	50 (14.2)	.85 [.75, .91]	.726	.498	1.594	1.245
Neglect responsibility	66 (18.7)	.65 [.52, .76]	.637	.207	0.864	1.328
Hazardous use	46 (13.0)	.58 [.42, .71]	.470	.411	0.708	1.906
Social/legal problems	22 (6.2)	.86 [.72, .99]	.630	.675	1.683	1.805
"Binge" drinking	158 (44.8)	.58 [.46, .68]	.700	-.147	0.711	0.226
Fit statistics						
-2 LL[a]		2,798.1	2,775.2			
AIC		-5,083.9	-5,084.8			
BIC		-10,205.0	-10,184.1			

Notes: Discrim. = discrimination. LL = log likelihood; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion. [a]Likelihood ratio test: $\chi^2(10) = 22.97$, $p = .01$.

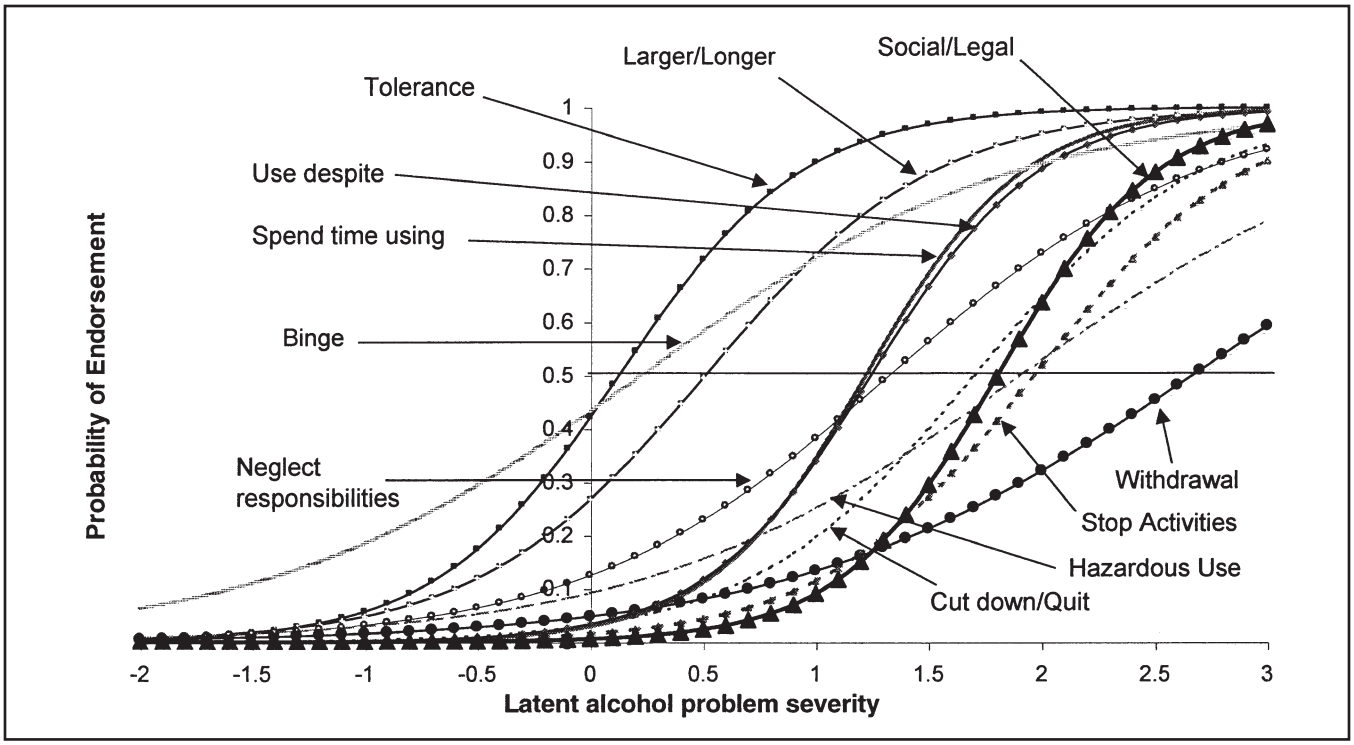

FIGURE 1. Item characteristic curves from left to right at 50% endorsement in 353 undergraduate students: tolerance, "binge" drinking, larger/longer, spend time using, use despite problems, neglect responsibilities, tried to quit or cut down on use, social/legal problems, hazardous use, stop activities to use, and withdrawal

the lowest (Table 1). At a 50% probability of endorsing an item, all values fell between 0.13 and 2.7 of the standardized severity continuum. Severity parameters for drinking despite problems, time spent drinking, and neglecting responsibilities to drink were nearly indistinguishable (Table 1 and Figure 1). Likewise, severity parameters for social/legal problems, hazardous use, activities reduced, and cut down/quit clustered together on the continuum. The severity parameters for withdrawal, hazardous use, and neglecting responsibilities were least able to distinguish levels of severity, and having social/legal problems, time spent drinking, and continued use despite problems were most able to distinguish levels of severity in this undergraduate sample. The parameters for such drinking fell in the least severe region of the item-characteristic curve showing both low severity and low discrimination. The ability of "binge" drinking to distinguish between severity levels was very similar to withdrawal, hazardous use, and neglecting responsibilities to drink.

Discussion

In this sample of undergraduates, a single latent severity factor best fit the DSM-IV AUD criteria. A two-factor model best fit the criteria including "binge" drinking, but no clearly defined second factor was identified. Past investigations have supported the robustness of IRT to violations of the unidimensionality assumption when there is support for one dominant dimension (Dorans and Kingston, 1985; Drasgow and Parsons, 1983; Harrison, 1986). The factor loading for "binge" drinking was greater than that for withdrawal and similar in magnitude to that for hazardous use. The three abuse criteria were interspersed among the seven dependence criteria in the IRT model. In accordance with prior findings (Gelhorn et al., 2008; Kahler and Strong, 2006; Martin et al., 2006; Saha et al., 2006), there was no evidence for a hierarchical structuring of abuse and dependence.

Notably, unlike prior investigations of the DSM-IV criteria, the heavy drinking item in this study included the 2-hour time qualifier advocated by the National Institute on Alcohol Abuse and Alcoholism (2004). Martin and colleagues (2008) argued that heavy drinking with the time qualifier should be included as an AUD criterion in DSM-V. As in other IRT analyses, the heavy drinking item in this study was found to be at the lowest severity level (Krueger et al., 2004; Nichol et al., 2007; Saha et al., 2007). If a measure of consumption should be added to cover the less severe region of the liability curve, weekly heavy episodic or "binge" drinking may satisfy this role. The prevalence of weekly "binge" drinking in this college sample was 43.8%; yet the sample was better at discriminating between levels of severity than withdrawal with a prevalence of 6.7%. "Binge" drinking showed factor loadings that were similar to other DSM-IV criteria, and IRT

analyses evidenced very low severity comparable to that for tolerance.

The ordering of criteria on the severity continuum differs from that in prior studies. An IRT of adolescents (Martin et al., 2006) found that tolerance had very low severity and withdrawal had very high severity—as we found, too, in our results—but drinking more than intended clustered with neglecting responsibilities and time spent drinking. In the present study, drinking more than intended showed greater severity than tolerance but did not cluster with other criteria. In the NESARC sample of current drinkers with current criteria, heavy episodic drinking also was shown to exhibit very low severity and landed near drinking larger and longer than intended, but tolerance was on the high severity end of the spectrum (Saha et al., 2006).

Limitations of this study should be noted. Social and legal problems were combined into one criterion because of concerns specific to the university where the data were collected. To keep alcohol-related self-reports other than the AUD items to the well-validated 3-month time frame often used in investigations of undergraduate drinking, the "binge"-drinking item was reported in a different time frame (the past 3 months) than the self-report items on the DSM-IV AUD criteria (past 12 months, in keeping with the DSM-IV). The high levels of such drinking made the present sample particularly appropriate for addressing the addition of this item to the DSM-IV AUD criteria.

Women were disproportionately represented in the present sample, and this gender disparity precluded running of the factor analyses and IRT analyses within each gender. Although fewer women than men in the present sample met the criterion for weekly "binge" drinking, they nonetheless tended to drink heavily (41.1% met such weekly drinking criterion), and men and women differed on only 1 of the 10 DSM-IV AUD items. Discrimination and severity parameters were not tested for statistically significant differences.

This is the first study to examine the DSM-IV criteria for AUDs in a college sample using IRT. In this sample, abuse, dependence, and "binge" drinking appeared to be intermixed on a continuum of severity. "Binge" drinking was in the least severe region of the continuum, falling between tolerance and drinking more than intended.

References

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.). Washington, DC: Author.

Dawson, D. A., Grant, B. F., Stinson, F. S., & Chou, P. S. (2004). Another look at heavy episodic drinking and alcohol use disorders among college and noncollege youth. *Journal of Studies on Alcohol, 65,* 477-488.

Dorans, N. J., & Kingston, N. M. (1985). The effects of violations of unidimensionality on the estimation of item and ability parameters and on item response theory equating of the GRE verbal scale. *Journal of Educational Measurement, 22,* 249-262.

Drasgow, F., & Parsons, C. K. (1983). Application of unidimensional item response theory models to multidimensional data. *Applied Psychological Measurement, 7,* 189-199.

Gelhorn, H., Hartman, C., Sakai, J., Stallings, M., Young, S., Rhee, S. H., Corley, R., Hewitt, J., Hopfer, C., & Crowley, T. (2008). Toward DSM-V: An item response theory analysis of the diagnostic process for DSM-IV alcohol abuse and dependence in adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 47,* 1329-1339.

Gillespie, N. A., Neale, M. C., Prescott, C. A., Aggen, S. H., & Kendler, K. S. (2007). Factor and item-response analysis DSM-IV criteria for abuse of and dependence on cannabis, cocaine, hallucinogens, sedatives, stimulants and opioids. *Addiction, 102,* 920-930.

Harrison, D. A. (1986). Robustness of IRT parameter estimation to violations of the unidimensionality assumption. *Journal of Educational Statistics, 11,* 91-115.

Hasin, D. S. & Beseler, C. L. (2009). Dimensionality of lifetime alcohol abuse, dependence and binge drinking. *Drug and Alcohol Dependence, 101,* 53-61.

Jennison, K. M. (2004). The short-term effects and unintended long-term consequences of binge drinking in college: A 10-year follow-up study. *American Journal of Drug and Alcohol Abuse, 30,* 659-684.

Johnston, L. D., O'Malley, P. M., & Bachman, J. G. (2003). Monitoring the future: National survey results on drug use, 1975-2002: Volume II. College students and adults ages 19-40. Bethesda, MD: National Institute on Drug Abuse.

Kahler, C. W., & Strong, D. R. (2006). A Rasch model analysis of DSM-IV alcohol abuse and dependence items in the National Epidemiological Survey on Alcohol and Related Conditions. *Alcoholism: Clinical and Experimental Research, 30,* 1165-1175.

Knight, J. R., Wechsler, H., Kuo, M., Seibring, M., Weitzman, E. R., & Schuckit, M. A. (2002). Alcohol abuse and dependence among U.S. college students. *Journal of Studies on Alcohol, 63,* 263-270.

Krueger, R. F., Nichol, P. E., Hicks, B. M., Markon K. E., Patrick, C. J., Iacono, W. G., & McGue, M. (2004). Using latent trait modeling to conceptualize an alcohol problems continuum. *Psychological Assessment, 16,* 107-119.

Langenbucher, J. W., Labouvie, E., Martin, C. S., Sanjuan, P. M., Bavly, L., Kirisci, L., & Chung, T. (2004). An application of item response theory analysis to alcohol, cannabis, and cocaine criteria in DSM-IV. *Journal of Abnormal Psychology, 113,* 72-80.

Li, T.-K., Hewitt, B. G., & Grant, B. F. (2007a). The Alcohol Dependence Syndrome, 30 years later: A commentary. *Addiction, 102,* 1522-1530.

Li, T.-K., Hewitt, B. G., & Grant, B. F. (2007b). Is there a future for quantifying drinking in the diagnosis, treatment, and prevention of alcohol use disorders? *Alcohol and Alcoholism, 41,* 57-63.

Martin, C. S., Chung, T., Kirisci, L., & Langenbucher, J. W. (2006). Item response theory analysis of diagnostic criteria for alcohol and cannabis use disorders in adolescents: Implications for DSM-V. *Journal of Abnormal Psychology, 115,* 807-814.

Martin, C. S., Chung, T., & Langenbucher, J. W. (2008). How should we revise diagnostic criteria for substance use disorders in the DSM-V? *Journal of Abnormal Psychology, 117,* 561-575.

National Institute on Alcohol Abuse and Alcoholism. (2004, winter). *NIAAA council approves definition of binge drinking* (NIAAA Newsletter, No. 3, NIH Publication No. 04-5346, p. 3). Bethesda, MD: Author. Retrieved from http://pubs.niaaa.nih.gov/publications/Newsletter/winter2004/Newsletter_Number3.htm#council

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: Statistical modeling* (6th ed.). Richmond, VA: Department of Psychiatry, Virginia Commonwealth University.

Nichol, P. E., Krueger, R. F., & Iacono, W. G. (2007). Investigating gender differences in alcohol problems: A latent trait modeling approach. *Alcoholism: Clinical and Experimental Research, 31,* 783-794.

O'Malley, P. M., & Johnston, L. D. (2002). Epidemiology of alcohol and other drug use among American college students. *Journal of Studies on Alcohol, Supplement No. 14,* 23-39.

Perkins, H. W. (2002). Surveying the damage: A review of research on consequences of alcohol misuse in college populations. *Journal of Studies on Alcohol, Supplement No. 14,* 91-100.

Proudfoot, H., Baillie, A. J., & Teesson, M. (2006). The structure of alcohol dependence in the community. *Drug and Alcohol Dependence, 81,* 21-26.

Saha, T. D., Chou, S. P., & Grant, B. F. (2006). Toward an alcohol use disorder continuum using item response theory: Results from the National Epidemiological Survey on Alcohol and Related Conditions. *Psychological Medicine, 36,* 931-941.

Saha, T. D., Stinson, F. S., & Grant, B. F. (2007). The role of alcohol consumption in future classifications of alcohol use disorders. *Drug and Alcohol Dependence, 89,* 82-92.

Slutske, W. S. (2005). Alcohol use disorders among US college students and their non-college-attending peers. *Archives of General Psychiatry, 62,* 321-327.

Slutske, W. S., Hunt-Carter, E. E., Nabors-Oberg, R. E., Sher, K. J., Bucholz, K. K., Madden, P. A., Anokhin, A., & Heath, A. C. (2004). Do college students drink more than their non-college-attending peers? Evidence from a population-based longitudinal female twin study. *Journal of Abnormal Psychology, 113,* 530-540.

Wechsler, H., Lee, J. E., Nelson, T. F., & Kuo, M. (2002). Underage college students' drinking behavior, access to alcohol, and the influence of deterrence policies: Findings from the Harvard School of Public Health College Alcohol Study. *Journal of American College Health, 50,* 223-236.

Wechsler, H., & Nelson, T. F. (2001). Binge drinking and the American college student: What's five drinks? *Psychology of Addictive Behaviors, 15,* 287-291.