



Published in final edited form as:

Drug Alcohol Depend. 2011 November 1; 118(2-3): 202–208. doi:10.1016/j.drugalcdep.2011.03.021.

Toward DSM-V: Mapping the Alcohol Use Disorder Continuum in College Students

Brett T. Hagman and

Department of Mental Health Law and Policy, University of South Florida, 13301 Bruce B. Downs Blvd, MHC 1701, Tampa, Florida 33612

Amy M. Cohn

Department of Mental Health Law and Policy, University of South Florida, 13301 Bruce B. Downs Blvd, MHC 2716, Tampa, Florida 33612

Abstract

Objective—The present study examined the dimensionality of DSM-IV Alcohol Use Disorder (AUD) criteria using Item Response Theory (IRT) methods and tested the validity of the proposed DSM-V AUD guidelines in a sample of college students.

Method—Participants were 396 college students who reported any alcohol use in the past 90 days and were aged 18 years or older. We conducted factor analyses to determine whether a one- or two-factor model provided a better fit to the AUD criteria. IRT analyses estimated item severity and discrimination parameters for each criterion. Multivariate analyses examined differences among the DSM-V diagnostic cut-off (AUD versus No AUD) and severity qualifiers (no diagnosis, moderate, severe) across several validating measures of alcohol use.

Results—A dominant single-factor model provided the best fit to the AUD criteria. IRT analyses indicated that abuse and dependence criteria were intermixed along the latent continuum. The "legal problems" criterion had the highest severity parameter and the tolerance criterion had the lowest severity parameter. The abuse criterion "social/interpersonal problems" and dependence criterion "activities to obtain alcohol" had the highest discrimination parameter estimates. Multivariate analysis indicated that the DSM-V cut-off point, and severity qualifier groups were distinguishable on several measures of alcohol consumption, drinking consequences, and drinking restraint.

Discussion—Findings suggest that the AUD criteria reflect a latent variable that represents a primary disorder and provide support for the proposed DSM-V AUD criteria in a sample of college students. Continued research in other high-risk samples of college students is needed.

Keywords

College students; Item Response Theory; reliability; validity; alcohol use

© 2011 Elsevier Ireland Ltd. All rights reserved.

Correspondence can be address to the first author at: Brett T. Hagman, Department of Mental Health Law and Policy, University of South Florida, 1330 Bruce B. Downs Blvd, MHC 1701, Tampa, Florida 33612, Phone: (813) 974-9004, Fax: (813) 974-9327, hagman@usf.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.

1. Introduction

The Diagnostic and Statistical Manual of Mental Disorders 4th edition (DSM-IV; APA, 2000) conceptualizes alcohol abuse and dependence as two separate factors with a hierarchical relationship (Kahler and Strong, 2006). Concerns have been raised over the utility of this categorical approach; primary among them is whether the AUD criteria should be combined into a single disorder that reflects a continuum of alcohol use severity. An additional concern is related to the sensitivity/specificity of classification (Pollock and Martin, 1999). For example, under DSM-IV, there exists a sub-set of individuals who endorse a sub-threshold number of symptoms of dependence, but none for abuse, and do not receive a formal diagnosis (i.e., “diagnostic orphans”), despite findings that these individuals have rates of consumption and problem severity that are comparable to those with abuse (Pollock and Martin, 1999). Thus, a percentage of at-risk individuals go undetected. These concerns have led to proposed revisions for the upcoming DSM-V (APA, 2010; Hasin and Beseler, 2009).

Research using population-based and clinical samples of adults and adolescents, but not college students, has shown mixed results for the categorical structure of the DSM-IV AUD criteria (Bucholz et al., 1996; Kahler and Strong, 2006; Lynskey et al., 2005; Martin et al., 2006; Saha et al., 2006). For example, test-retest reliability studies have shown moderate to high reliability estimates for alcohol dependence but lower reliability estimates for alcohol abuse (Hasin et al., 2006; Horton et al., 2000). Results from factor analytic studies of the AUD criteria are equivocal. Some indicate better fit for a two factor model that supports the current diagnostic system, but also show high factor correlations suggesting substantial overlap (Harford and Muthen, 2000; Grant et al., 2007); whereas others show support for a dominant single factor (Martin et al., 2006; Saha et al., 2006). Overall, findings from test-retest and factor analytic studies provide inconclusive evidence with respect to the dimensionality of AUD criteria.

Item response theory (IRT) methods can improve upon prior statistical assessment techniques by providing information on rank order of each criterion in relation to overall diagnostic severity. *Item location* (severity) parameters provide information on the severity of an item (Embretson and Reise, 2000), whereas *Item discrimination* (slope) parameters provide information on the strength of an item in relation to the underlying latent trait. Initial research using IRT methods to examine the dimensionality of DSM-IV AUD criteria in community and clinical samples indicate that the criteria are intermixed along a latent alcohol use severity continuum (Dawson et al., 2010; Gelhorn et al., 2008; Kahler and Strong, 2006; Martin et al., 2006; Saha et al., 2006; Shmulewitz et al., 2010), suggesting no defined rank order among the criteria and that the categorical distinction may not be supported. Only one study to our knowledge has used IRT to assess the dimensionality of DSM-IV AUD criteria in a college student sample (Beseler et al., 2010). Consistent with prior IRT studies, Beseler and colleagues (2010) showed that the criteria were intermixed along an AUD severity continuum, which provides good preliminary evidence of a non-hierarchical structure of the AUD criteria in college students. However, the sample was predominantly female and White, the limitations of which suggest that further replication of the IRT findings should be warranted with a more heterogeneous sample.

In response to evidence that the DSM-IV AUD criteria do not represent distinct diagnostic entities, the DSM-V Substance Use Disorders task force has recommended that the new criteria be combined into one single disorder (APA, 2010), with “severity” indicators such that individuals who endorse at least 2 of any criteria will receive an AUD diagnosis, while those who endorse between 2 to 3 criteria or 4 or more criteria will receive an AUD diagnosis of “moderate” or “severe”, respectively. No research has, however, examined the

validity of the proposed DSM-V modifications with a college student sample, despite the fact that half of all college students engage in heavy drinking (O'Malley and Johnston, 2002; Wechsler et al., 2002), and between 6 and 31% meet diagnostic criteria for past-year alcohol abuse or dependence (Dawson et al., 2004; Knight et al., 2002).

These proposed DSM-V guidelines offer a unique opportunity to provide empirical support for these recommendations. Given that only one study has employed IRT with a college student sample to examine the dimensionality of DSM-IV AUD criteria, and no study has examined the applicability of the proposed DSM-V criteria with a college student sample, the current study aims were to 1) partially replicate prior findings (Beseler, 2010) by examining the dimensionality of DSM-IV AUD criteria in a sample of college students using IRT, and 2) evaluate the convergent validity of the proposed DSM-V guidelines.

2. Methods

2.1 Participants and Procedures

The sample (N = 496) was collected at a large public northeastern university in the United States. Enrollment occurred throughout the spring 2010 academic semester. Most (N = 457) were recruited from Introductory Psychology courses and received course credit, and a minority (n = 39) were recruited via newspaper advertisements, and paid \$15.00. Recruitment techniques did not differ on select demographic variables or relevant alcohol use variables. All participants provided informed consent and completed a one-hour anonymous battery of questionnaires. All procedures were approved by the university's Institutional Review Board.

Participants were eligible if they were 18 years of age or older, an undergraduate student, and consumed any alcohol in the prior 90 days. Non-drinkers were excluded from analyses, leaving a total of 396 participants for the IRT analyses. The sample of 396 (Mean age = 19, SD = 1.28) was 52% male and racially/ethnically diverse: 58.8% (n = 237) Caucasian, 19.9% (n = 79) Asian, 8.3% (n = 33) African-American, 6.6% (n = 26) Hispanic, and 5.3% (n = 21) labeled as "other." Most were full-time students (66.4%; n = 263), and lived on campus (78.5%; n = 300). With respect to class rank, 59.3% (n = 235) were freshmen, 20.2% (n = 80) sophomores, 14.1% (n = 56) juniors, and 6.4% (n = 25) seniors.

2.2 Measures

2.2.1 DSM-IV Alcohol Use Disorder Criteria—A total of 13 questions were created from the 11 DSM-IV alcohol abuse/dependence criteria to determine the diagnosis of an AUD (APA, 2000), and paralleled the alcohol use diagnostic questions from the Composite International Diagnostic Interview-Substance Abuse Module (CIDI-SAM; Clements, 1998; Robins et al., 1990). Participants were asked to report on the occurrence (yes) or absence (no) of each criterion within the past year.

2.2.2 Alcohol and Drug Use—Alcohol and drug use in the prior 90-days were collected via a modified quantity/frequency index (QFI; see Hagman et al., 2007). Respondents estimated their frequency of consuming hard liquor, wine, and beer (*1 = never; 7 = almost everyday*), and the quantity of alcohol they consumed per drinking occasion in the following three categories: hard liquor (*1 = never; 7 = 16 or more shots*), wine (*1 = never; 7 = 16 or more 5-oz glasses of wine*), and beer (*1 = never; 7 = 16 or more 12-oz cans/bottles*). Separate frequency and quantity indices were created by summing reports for beer, wine and liquor items. Principal components analyses indicated that each index had a uni-dimensional structure and accounted for 51.4% and 47.6% of the common variance for frequency and quantity indices, respectively.

Additionally, participants were asked to report the “typical” number of standard drinks consumed per weekday and weekend, the largest amount of alcohol consumed (in standard drinks) in a 24-hour period, and frequency (1 = *never*; 7 = *almost every day*) of using marijuana, hashish, crack, cocaine, amphetamines, barbiturates, benzodiazepines, opiates, and hallucinogens.

2.2.3 College Alcohol Problem Scale-Revised (CAPS-r)—The CAPS-r (8 items; Maddock et al., 2009) measured social and personal consequences (1 = *never* to 6 = *10 or more times*) associated with alcohol use in the past 12-months. Items were dichotomized (0 = *did not occur* to 1 = *did occur*) and summed together to create an index of problem severity ($\alpha = .70$).

2.2.4 Temptation and Restraint Inventory (TRI)—The TRI (15-item; Collins and Lapp, 1992) measured temptation to drink alcohol and attempts to control alcohol use (1 = *Never* to 9 = *Always*). The three lower-order factors of Govern (difficulty controlling drinking; $\alpha = .82$), Emotion (drinking for negative emotions; $\alpha = .81$), and Preoccupation (thoughts about drinking; $\alpha = .77$) were used in the analyses.

2.3 Data Analytic Plan

2.3.1 IRT and Factor Analysis—Exploratory (EFA) and confirmatory factor analyses (CFA) were conducted on the 11 criteria to ensure that IRT assumptions were met. For the EFA, Catell’s Scree Plot (Catell, 1966) and the Kaiser criterion (i.e., eigenvalues > 1) were used to determine the number of factors to retain. The overall adequacy of model fit was determined using multiple indices (Hu and Bentler, 1995, 1999). We used MPlus 6.0 software to conduct all factor analyses, which uses a robust unweighted least squares estimation for the EFA and weighted least squares estimation for the CFA procedure to derive estimates (Muthen and Muthen, 1998). Tetrachoric correlations were specified.

Two-parameter IRT models were applied to the 11 abuse and dependence criteria, for estimates of item *location* or severity parameters and item *discrimination* or slope parameters. Item misfit analysis was then conducted to determine whether each criterion fit the specified two-parameter models. Final IRT analyses were then conducted and item characteristic curves (ICCs) were plotted. The ICCs provide a graphical depiction of the probability that a specific item is endorsed as a function of the value of the purported underlying latent-trait continuum (Embretson and Reise, 2000). All IRT models were analyzed using Parscale 4.1 (Scientific Software International, 2003), which estimates item parameters via a Bayesian expectation-maximization equation. The convergence criterion was set to .001.

2.3.2 Evaluating Validity of the DSM-V Guidelines—Prevalence estimates of the current DSM-IV alcohol abuse and dependence criteria were compared to the proposed DSM-V criteria. We examined discrepancies in classification status and descriptive differences across several external validating measures of alcohol use, drinking restraint, and problem severity. For the classification of the DSM-IV AUD diagnoses, cases were defined as “No AUD diagnosis” for participants with no criteria endorsement, “Abuse” for participants who endorsed ≤ 4 abuse criteria and < 3 dependence criteria, and “Dependence” for participants who endorsed ≥ 3 dependence criteria. We also included a fourth category of “Diagnostic Orphans” to capture individuals who endorsed ≤ 2 dependence and no abuse criteria and did not meet DSM-IV criteria (Pollock and Martin, 1999). We also examined the DSM-V AUD diagnostic algorithm (APA, 2010), by comparing those with “no AUD diagnosis” (≤ 1 criteria) to those with an “AUD diagnosis” (≥ 2 criteria).

Second, using MANOVA and Hotellings T^2 tests, we examined differences between the DSM-IV subgroups (no-AUD diagnosis, diagnostic orphans, abuse diagnosis, dependence diagnosis) and proposed DSM-V diagnostic subgroups (No AUD diagnosis vs. AUD diagnosis) across measures of alcohol consumption (quantity and frequency), drinking restraint, and problem severity. Tukey's post-hoc test for planned comparisons was used to control for Type 1 error. All results were considered significant at the $p < .01$ level.

Lastly, we conducted a MANOVA test to examine differences across the AUD severity groups (i.e., no diagnosis, moderate, severe) on measures of alcohol consumption, drinking restraint, and problem severity. Each group represented increasing severity of alcohol problems with those who endorsed ≤ 1 AUD criteria classified as "no diagnosis", > 1 or ≤ 3 AUD criteria classified as "moderate", and ≥ 4 criteria classified as "severe". Tukey's post-hoc test for planned comparisons was used to control for Type 1 error. All results were considered significant at the $p < .01$ level.

4. Results

4.1 Drinking and Drug Use Characteristics of the Sample

Participants drank, on average, 17 out of the prior 90 days (M # drinking days = 17.29; SD = 14.46), and consumed 2.54 (SD = 2.73) standard drinks per day on the weekday and 5.84 (SD = 3.93) standard drinks per day on the weekend. On average, the greatest amount of alcohol consumed within a 24-hour period was 9.67 standard drinks (SD = 5.68), with a greater number of participants reporting weekly consumption of beer (48.1%) than either liquor (42.1%) or wine (6.3%). Forty-seven percent of the sample reported marijuana use in the past 90-days, while the frequency of other illicit drug use was low.

3.2 Factor Analyses: Assessment of IRT Model Assumptions

Results from the EFA and CFA indicated that the 11 AUD criteria reflected a uni-dimensional construct with a strong dominant first higher-order factor (Table 1). Eigenvalues for the first and second factors were 4.191 and 1.474, respectively. The correlation between factors was in the moderate range ($r = .47$) indicating overlap between factors. The first factor accounted for 38.10% of the variance, in comparison to the second factor accounting for 13.40% of the variance. A CFA indicated good model fit for a single dimensional structure of the 11 AUD criteria: TLI = .93, CFI = .95, and RMSEA = .04.

3.3 Item Misfit Analysis

The two-parameter IRT models fit the majority of the criteria, but did not provide a good fit for "consumed alcohol over larger/longer period than intended (dependence)" χ^2 (6, N = 392) = 60.02, $p < .001$, and tolerance (dependence) χ^2 (6, N = 392) = 80.19, $p < .01$. The chi-square fit test for the overall model was not significant, indicating good overall model fit.

3.4 Final IRT Model Analysis

Endorsement for each of the AUD criteria ranged from 4.5% to 50%. The items "consumed alcohol over larger/longer period than intended" (Tolerance/dependence)" and social/interpersonal problems" (abuse) had the highest levels of endorsement, whereas withdrawal (dependence), "legal problems" (abuse), "important activities given up" (dependence), and "physically hazardous" (dependence) had the lowest frequency of endorsement. As shown in Table 1, severity (location) parameters ranged from .018 to 3.841 with an overall mean of 2.047(SE = 1.185) across the continuum. The lowest severity parameters were found for tolerance (dependence) and "consumed alcohol over larger/longer period than intended" (dependence). This implies that these items were less severe and more likely to be endorsed.

The severity parameters for “legal problems” (abuse), and withdrawal (dependence) were higher, indicating that these items were more severe and less likely to be endorsed by the sample. The remaining criteria were intermixed along the AUD severity continuum, and were located between the highest and lowest severity parameters.

The overall mean for the discrimination parameters was 0.754 (SE = 0.285) and ranged from .431 to 1.241. The lowest parameters were found for “legal problems (abuse)”, “role obligations (abuse)”, “physically hazardous (abuse)”, tolerance (dependence) and withdrawal (dependence) indicating that these items did not provide a high degree of discrimination across the latent alcohol use problem severity continuum. Those with the highest discrimination parameters were “unsuccessful efforts (dependence)”, “recurrent physical/psychological problems (dependence)”, “social/interpersonal problems (abuse)”, and “important activities given up (dependence)”. See Figure 1 for ICCs.

3.5 Descriptive Analysis of Proposed DSM-V Cut-Off

DSM-IV rates of abuse and dependence were 23.2% (n = 92) and 16.7% (n = 66), respectively; with 39.9% (n = 158) who met diagnostic criteria for any AUD under the DSM-IV system (Table 2). Notably, 34.1% (n = 135) were classified as “diagnostic orphans”, but experienced a number of alcohol-related negative consequences on the CAPS-r and had high rates of alcohol consumption (Table 3). MANOVA results comparing differences among groups were significant [Wilks’ Lambda = .514, $F(18, 1089) = 16.07$, $p < .001$]. All follow-up univariate F-tests were significant (p ’s < .001). Post-hoc analyses showed that those with a dependence diagnosis were more severe on all indicators ($p < .01$) relative to the other three DSM-IV AUD groups. Those with an abuse diagnosis and “diagnostic orphans” were more severe than those with no diagnosis, but not significantly different from each other (Table 3).

Approximately half (n = 198) of the sample met criteria for an AUD diagnosis using the DSM-V algorithm (Table 4). Further, 18.18% of participants who met DSM-V criteria for an AUD met DSM-IV criteria for abuse (n = 72) and 16.7% (n = 66) for dependence, while 5% (n = 20) who met criteria for DSM-IV abuse did not meet criteria for a DSM-V AUD diagnosis. Lastly, 15.15 % (n = 60) of those who were classified as “diagnostic orphans” under the DSM-IV algorithm met diagnostic criteria for a DSM-V AUD diagnosis.

3.6 Convergent Validity of DSM-V AUD criteria

Table 4 displays results of the Hotellings T^2 that examined mean differences between the proposed DSM-V diagnostic groups (i.e., No AUD diagnosis vs. AUD diagnosis) across the external validator variables of alcohol use, drinking restraint, and problem severity. The overall omnibus test was significant [Wilks’ Lambda = .697, $F(6, 387) = 28.09$, $p = .001$] and all follow-up univariate t-tests were significant (all p ’s < .01). Compared to those who did not meet criteria for a DSM-V alcohol use disorder (i.e., No AUD diagnosis), those with an AUD diagnosis reported greater alcohol-related negative consequences, drank alcohol more frequently and in greater quantities, and had higher levels of drinking restraint.

There were nearly three times as many participants classified as “moderate” severity (36.16%) than “severe” (13.38%). Reports of alcohol use consumption, drinking restraint, and problem severity among those in the “moderate” and “severe” groups were similar to the DSM-IV abuse and dependence diagnoses, respectively (Table 5). Results of the MANOVA2 between the AUD severity qualifier groups was significant [Wilks’ Lambda = .557, $F(12, 772) = 20.38$, $p < .01$], and all univariate F-tests were statistically significant (all p ’s < .01). Post-hoc analyses indicated that the three groups were significantly different from

one another (all p 's < .01), with those in the “severe” group reporting the highest levels on each of the external validators compared to the other severity qualifier groups.

4. Discussion

The first aim of this study was to examine the dimensionality of the DSM-IV AUD criteria. Exploratory and confirmatory factor analyses indicated that a single latent alcohol use severity factor provided the best fit. These findings are consistent with prior research supporting a dominant single factor of the AUD criteria (Beseler et al., 2010; Gelhorn et al., 2008; Saha, et al., 2006). Further, and consistent with prior studies (Beseler et al., 2010; Gelhorn et al., 2008; Saha et al., 2006), our IRT results indicated that the abuse criteria were intermixed with the dependence criteria across the latent alcohol use disorder severity continuum, and therefore did not support a hierarchical ordering; while the location parameters in this study were plotted in the middle to high end of the alcohol problem severity continuum, suggesting these items are most sensitive at classifying college students with higher levels of severity. We recommend combining the criteria to reflect a single AUD continuum as indicated by the DSM-V Substance Use Task Force (APA, 2010).

Some inconsistencies between our findings and those reported in prior IRT studies bear further consideration. First, and consistent with reports by Beseler et al. (2010), a closer inspection of the AUD criteria in this study indicated that tolerance (dependence) and “drinking in larger/longer amounts than intended (dependence)” fell at the lower end of the continuum, while withdrawal (dependence), “legal problems (abuse)” and “failure to fulfill role obligations (abuse)” fell at the higher end. Beseler et al., (2010) also found that tolerance and “drinking in larger/longer amounts than intended” mapped onto the lower end of the continuum, while withdrawal fell onto the higher end of the spectrum. Conversely, our IRT findings indicate that the criteria “recurrent physical/psychological problems (dependence)” and “role obligations (abuse)” were more severe than those reported in the Beseler et al., (2010) study. Further, Saha and colleagues (2006), who used a population-based sample aged 18 to 65, showed that the criteria “important activities given up (dependence)” and “role obligations (abuse)” had the highest location (severity) parameter estimates, which is not consistent with our findings. Interestingly, Saha et al (2006) removed the “legal problems (abuse)” criterion because of poor model fit, whereas this item was included in our analyses. Differences in the placement of specific location (severity) parameters across studies may likely reflect demographic differences, differences in the methods used to obtain diagnostic information, the separation of the abuse criteria “legal problems” and “social/interpersonal problems”, or the use of different statistical packages to derive parameter estimates (Childs and Chen, 1999).¹ More research is needed using diverse

²We recognize that several different severity qualifier groups are still under active consideration by the DSM-V Substance Use Task Force. To address this, we also examined differences among undiagnosed, mild, moderate, severe individuals using the following alternative severity group indicators: Those who endorsed ≤ 1 AUD criteria were classified as “no diagnosis”, ≥ 2 or ≤ 3 AUD criteria classified as “mild”, ≥ 4 or ≤ 5 criteria classified as “moderate,” and ≥ 6 criteria classified as “severe”. Overall, fifty percent of the sample ($n = 198$) were classified as “no diagnosis”, 36.61% ($n = 145$) classified as “minimal” severity, 9.1% ($n = 36$) classified as “moderate” severity, and 4.29% ($n = 17$) classified as “severe”. We conducted a MANOVA test to examine differences across these AUD severity groups on measures of alcohol consumption, drinking restraint, and problem severity. Results were significant [Wilks' Lambda = .544, $F(18, 1095) = 14.59, p < .01$], and all univariate F-tests were statistically significant (all p 's < .01). Post-hoc analyses indicated those in the “moderate” and “severe” groups were not significantly different from one another. However, the “moderate” and “severe” qualifier groups were significantly more severe on all outcome variables from those in the “mild” and the “no diagnosis” qualifier groups; and those in the “mild” group were significantly more severe from those in the “no diagnosis” group. While these findings do not provide strong support for an alternative severity qualifier indicator (no diagnosis, minimal, moderate, severe) in a sample of college students, they should be interpreted with caution given the small sample size in the “severe” qualifier group.

¹In order to understand more thoroughly the differences in the IRT severity parameters obtained in our study from those presented in the Beseler et al., (2010) study, we examined differential item functioning (DIF) across each of the 11 DSM-IV criteria in the following groups: 1) gender (male vs. female); and 2) illicit drug use status (no illicit drug use vs. any illicit drug use). Results indicated no significant DIF for each of the 11 DSM-IV AUD criteria by gender or illicit drug use status. Results may be obtained from the first author (Brett T. Hagman).

samples of college students to understand the stability of the IRT location (severity) parameters.

A second aim of this study was to examine the validity of the proposed DSM-V AUD criteria in a sample of college students. Descriptive analyses indicated that a greater number of individuals met criteria for an AUD under the proposed DSM-V diagnostic system compared to those who met criteria for an AUD under the current DSM-IV system. Further, a small group of cases (5%) who met DSM-IV criteria for abuse did not meet criteria for a DSM-V AUD. This suggests that some individuals will not receive a diagnosis under the new DSM-V diagnostic system, despite being at risk for alcohol-related problems. This highlights a potential disadvantage to the new proposed system. Another novel finding of this study concerns the relatively high number of individuals classified as “diagnostic orphans” under the current DSM-IV system (34.1%) who did receive a diagnosis under DSM-V guidelines (15%). This finding suggests that a substantial proportion of “sub-threshold” individuals will not go undetected under the newly proposed system, which is an improvement over DSM-IV. This latter finding indicates an important discrepancy between each classification system with respect to sensitivity and specificity; more importantly, neither system appears to be entirely effective at diagnosing individuals with lower levels of problem severity.

Overall, findings do support the convergent validity of the proposed DSM-V diagnostic system in college students, as the groups can be distinguished between each other on several external validating measures of alcohol use, drinking restraint and problem severity. Multivariate analyses indicated that those who met DSM-V criteria for an AUD were more severe on alcohol consumption variables, and related correlates compared to those who did not meet DSM-V criteria, while those in the “moderate” and “severe” groups had mean rates of alcohol consumption, drinking restraint and alcohol problems that were equivalent to those classified with abuse and dependence under the DSM-IV guidelines, respectively. As would be expected, those in the “severe” DSM-V AUD group had the highest levels of alcohol consumption, drinking restraint, and problem severity compared to those in the “moderate” and “no/minimal” severity groups. Continued research is needed to validate the DSM-V diagnostic threshold and severity qualifiers among other high-risk samples of drinkers in college, and compare to same-aged non-college student peers.

There were some limitations of the current study. First, the sample was primarily recruited through an undergraduate research pool, which may limit generalizability. However, drinking rates in the current sample are similar to those reported in other epidemiological studies of college student drinking (O’Malley and Johnston, 2002). Second, we used self-report assessment of AUD criteria rather than a semi-structured diagnostic interview. However, the items used to assess AUD symptoms in the current study were adapted from the CIDI-SAM (Robins et al., 1990), and mimic the wording found in the DSM-IV manual (APA, 2000). The validity of such assessments has been determined in large epidemiological studies, where it is not feasible to interview participants’ face-to-face (Dawson et al., 2004). Third, we did not examine differential item functioning across class rank due to disproportionate cell sizes. Future research should examine whether the AUD criteria perform similarly across different classes of college students. Lastly, it would be noteworthy to examine the validity of alternative algorithms for calculating severity qualifier thresholds using a larger sample size. Unfortunately, the current study sample size precludes us from making robust comparisons across groups of individuals who endorsed more than 5 AUD criteria.

There are several important conclusions from this study. First, IRT analyses indicated that alcohol abuse and dependence criteria are intermixed along a continuum of problem

severity, suggesting they should be combined to reflect a single disorder. Second, the proposed DSM-V diagnostic threshold is valid in this sample, and may assist in eliminating “diagnostic orphans” who do not receive a diagnosis. Alternatively, a minority of individuals who receive a DSM-IV abuse diagnosis do not received a diagnosis under the DSM-V system, indicating a limitation of the new diagnostic rules. Third, the DSM-V severity qualifier groups appear valid in this sample and correspond to the equivalent levels of alcohol use severity when compared to those classified with abuse or dependence under the DSM-IV. Finally, despite intensive intervention efforts across many college campuses, rates of AUD diagnoses within both DSM diagnostic systems remain high and warrant attention.

References

- American Psychiatric Association (APA). fourth edition. Washington, DC: 2000. Diagnostic and statistical manual of mental disorders. text revision (DSM-IV-TR)
- American Psychiatric Association (APA). DSM-5 development: substance-related disorders. Arlington, VA: 2010. on, Retrieved from <http://www.dsm5.org/ProposedRevisions/Pages/Substance-RelatedDisorders.aspx>
- 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.* 2010; 71:418–423.
- Buchholz KK, Heath AC, Reich T, Hesselbrock VM, Kramer JR, Numberger JI, Schuckit MA. Can we subtype alcoholism? A latent class analysis of data from relatives of alcoholics in a multicenter family study of alcoholism. *Alcohol. Clin. Exp. Res.* 1996; 20:1462–1471. [PubMed: 8947326]
- Cattell RB. The scree test for the number of factors. *Multivariate Behav. Res.* 1966; 1:245–276.
- Childs RA, Chen WH. Software note: Obtaining comparable item parameter estimates in MULTILOG and PARSCALE for two polytomous IRT models. *Appl. Psychol. Measure.* 23:371–379.
- Clements R. A critical evaluation of several alcohol screening instruments using the CIDI-SAM as a criterion measure. *Alcohol. Clin. Exp. Res.* 1998; 22:985–993. [PubMed: 9726267]
- Collins RL, Lapp WM. The temptation and restraint inventory for measuring drinking restraint. *Br. J. Addict.* 1992; 87:625–633. [PubMed: 1591514]
- Dawson DA, Grant BF, Stinson FS, Chou PS. Another look at heavy episodic drinking and alcohol use disorders among college and non-college youth. *J. Stud. Alcohol.* 2004; 65:477–488. [PubMed: 15378804]
- 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]
- Embretson, SE.; Reise, S. *Item Response Theory for Psychologists.* Mahwah, NJ: Erlbaum Publishers; 2000.
- Grant BF, Harford TC, Muthen BO, Yi H, Hasin DS, Stinson FS. DSM-IV alcohol dependence and abuse: further evidence of validity in the general population. *Drug Alcohol Depend.* 2007; 86:154–166. [PubMed: 16814489]
- 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]
- Hagman BT, Noel NE, Clifford PR. Social norms theory-based interventions: testing the feasibility of purported mechanism of action. *J. Am. Coll. Health.* 2007; 56:293–298. [PubMed: 18089512]
- Harford TC, Muthen BO. The dimensionality of alcohol abuse and dependence: a multivariate analysis of DSM-IV symptom items in the national longitudinal survey of youth. *J. Stud. Alcohol.* 2000; 62:150–157. [PubMed: 11327181]
- Hasin DG, Beseler CL. Dimensionality of lifetime alcohol abuse, dependence and binge drinking. *Drug Alcohol Depend.* 2009; 101:53–61. [PubMed: 19095379]

- Hasin D, Hatzenbuehler ML, Keyes K, Ogburn E. Substance use disorders: diagnostic and statistical manual of mental disorders, fourth edition (DSM-IV) and international classification of diseases, tenth edition (ICD-10). *Addiction*. 2006; 101 Suppl. 1:59–75. [PubMed: 16930162]
- Horton J, Compton W, Cottler LB. Reliability of substance user diagnoses among African-Americans and caucasians. *Drug Alcohol Depend*. 1998; 57:203–209. [PubMed: 10661671]
- Hu, L-T.; Bentler, P. Evaluating model fit. In: Hoyle, RH., editor. *Structural Equation Modeling. Concepts, Issues, and Applications*. London: Sage; 1995. p. 76-99.
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Modeling*. 1999; 6:1–55.
- Kahler CW, Strong DR. A rasch model analysis of DSM-IV alcohol abuse and dependence items in the national epidemiological survey on alcohol and related conditions. *Alcohol. Clin. Exp. Res*. 2006; 30:1165–1175. [PubMed: 16792564]
- Knight JR, Wechsler H, Kuo M, Seibring M, Weitzman ER, Schuckit MA. Alcohol abuse and dependence among U.S. college students. *J. Stud. Alcohol*. 2002; 63:263–270. [PubMed: 12086126]
- Lynskey MT, Nelson EC, Neuman RJ, Bucholz KK, Madden PF, Knopik VS, Slutske W, Whitfield JB, Martin NG, Heath AC. Limitations of DSM-IV operationalizations of alcohol abuse and dependence in a sample of Australian twins. *Twin Res. Hum. Gen*. 2005; 8:574–558.
- Maddock JE, Laforge RG, Rossi JS, O’Hare T. The college alcohol problems scale. *Addict. Behav*. 2001; 26:385–398. [PubMed: 11436930]
- 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]
- Muthen, BO.; Muthen, L. *Mplus*. Los Angeles: Muthen and Muthen; 1998.
- Neal DJ, Corbin R, Fromme K. Measurement of alcohol-related consequences among high school and college students: application of item response models to the rutgers alcohol problem index. *Psychol. Assess*. 2006; 18:402–414. [PubMed: 17154761]
- O’Malley PM, Johnston LD. Epidemiology of alcohol and other drug use among American college students. *J. Stud. Alcohol Suppl*. No. 2002; 14:23–39.
- Pollock NK, Martin CS. Diagnostic orphans: adolescents with alcohol symptoms who do not qualify for DSM-IV abuse or dependence diagnoses. *Am. J. Psych*. 1999; 156:897–901.
- Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke JD, Farmer A, Jablenski A, Pickens R, Reiger DA, Sartorius N, Towle LH. The Composite international diagnostic interview (CIDI): an epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch. Gen. Psychiatry*. 1988; 45:1069–1077. [PubMed: 2848472]
- Saha TD, Chou PS, 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]
- Scientific Software International, I. *Parscale 4.1*. Lincolnwood, IL: 2003.
- Sheffield FD, Darkes J, Del Boca FK, Goldman MS. Binge drinking and alcohol-related problems among community college students: implications for prevention policy. *J. Am. Coll. Health*. 2005; 54:137–141. [PubMed: 16335480]
- Shmulewitz D, Keyes K, Beseler C, Aharonovich E, Aivadyan C, Spivak B, Hasin D. The dimensionality of alcohol use disorders: results from Israel. *Drug Alcohol Depend*. 2010; 111:146–154. [PubMed: 20537809]
- Ullman SE. Mental health services seeking in sexual assault victims. *Women Ther*. 2007; 30:61–84.
- Wechsler H, Lee JE, Kuo M, Seibring M, Nelson TF, Lee H. Trends in college binge drinking during a period of increased prevention efforts. *J. Am. Coll. Health*. 2002; 50:203–217. [PubMed: 11990979]

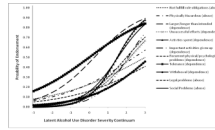


Figure 1.

Plot of item characteristic curves (ICC) for AUD criteria

Note: The X-axis represents the latent alcohol use problem severity continuum, and the Y-axis represents the probability of endorsement of each AUD criteria

Table 1
Final results of confirmatory factor analysis and two-parameter IRT analyses for the 10 AUD criteria

	% Endorsed	EFA Loadings (F1)	EFA Loadings (F2)	CFA Loadings	Location (SE)	Slope (SE)
Tolerance (dependence)	50%	0.407	0.426	0.473	0.018(.133)	0.552(.116)
Larger amounts than intended (dependence)	43.7%	0.544	0.445	0.583	0.288(.111)	0.745(.147)
Social/Interpersonal Problems (abuse)	12.9%	0.733	0.373	0.703	1.533(.214)	1.097(.242)
Activities to obtain alcohol (dependence)	11.9%	0.654	0.557	0.767	1.555(.187)	1.241(.302)
Recurrent physical/psychological problems (dependence)	10.1%	0.611	0.406	0.661	1.976(.306)	0.899(.235)
Important activities given up (dependence)	6.1%	0.689	0.411	0.715	2.199(.319)	1.074(.316)
Unsuccessful efforts (dependence)	8.1%	0.835	0.283	0.528	2.464(.423)	0.715(.288)
Physically hazardous (abuse)	7.9%	0.528	0.073	0.446	2.302(.515)	0.525(.155)
Role obligations (abuse)	9.9%	0.473	0.122	0.438	2.769(.927)	0.437(.126)
Withdrawal (dependence)	4.5%	0.650	-0.188	0.482	3.569(1.007)	0.575(.229)
Legal problems (abuse)	6.1%	0.539	0.076	0.472	3.841(1.005)	0.431(.241)
Eigenvalues		4.191	1.474			
% Variance accounted for		38.10%	13.40%			
Overall Mean (SE)					2.047(1.185)	0.754(.285)

Note: All values are sorted in ascending sequence according to the location (severity) parameters

% Endorsed reflects the overall percent endorsement for each criterion

EFA = Exploratory Factor Analysis; CFA = Confirmatory Factor Analysis

Table 2

Descriptive statistics (means, standard deviations) of AUD criteria endorsed between DSM-IV algorithm and proposed DSM-V algorithm

DSM-IV Diagnostic Category	n	Abuse Symptoms Endorsed		Dependence Symptoms Endorsed		AUD criteria endorsed		Sum of IRT severity parameters	
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
No AUD diagnosis	103(26%)	0	0	0	0	0	0	0	0
Diagnostic orphans	135(34.1%)	0	1.44(.49)	1.44(.49)	1.44(.49)	1.44(.49)	1.44(.49)	1.44(.49)	.50(.85)
Abuse diagnosis	92(23.2%)	1.28(.49)	1.09(.83)	1.09(.83)	2.38(.92)	2.38(.92)	1.56(3.41)	1.56(3.41)	1.56(3.41)
Dependence diagnosis	66(16.7%)	1.02(1.13)	3.55(.83)	3.55(.83)	4.56(1.61)	4.56(1.61)	5.73(3.35)	5.73(3.35)	5.73(3.35)
DSM-V Diagnostic Category	n	Abuse Symptoms Endorsed		Dependence Symptoms Endorsed		AUD criteria endorsed		Sum of IRT severity parameters	
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
No AUD diagnosis	198(50%)	.10(.30)	.38(.49)	.38(.49)	.48(.50)	.48(.50)	.34(.84)	.34(.84)	.34(.84)
DSM-V AUD diagnosis	198(50%)	.83(.93)	2.29(1.11)	2.29(1.11)	3.13(1.46)	3.13(1.46)	3.49(2.97)	3.49(2.97)	3.49(2.97)

Table 3
Descriptive statistics (means, standard deviations) of external validators by DSM-IV AUD diagnosis

DSM-IV AUD diagnosis	Overall		N = 103		N = 135		N = 92		N = 66	
	Mean(SD)	No Diagnosis(1)	Diagnostic Orphan(2)	Abuse diagnosis(3)	Dependence diagnosis(4)					
External validator variables	Mean(SD)	Mean(SD)	Mean(SD)	Mean(SD)	Mean(SD)					
Frequency of Alcohol Use Index	8.46(2.21)	6.94(2.07)	8.66(1.93)	8.95(2.01)	9.77(1.96)*					
Quantity of Alcohol Use Index	7.81(2.24)	6.34(1.95)	7.91(2.02)	8.34(1.96)	9.29(2.29)*					
CAPS-r sum scores	2.42(1.96)	1.15(1.44)	2.42(1.48)	3.08(1.91)	4.36(1.81)*					
Govern (TRI)	6.57(4.52)	4.09(2.23)	5.96(3.57)	6.66(3.74)	11.57(5.84)*					
Emotion (TRI)	7.06(4.25)	5.33(3.36)	6.24(3.13)	7.47(4.05)	10.84(5.30)*					
Preoccupation (TRI)	4.48(2.85)	3.45(1.48)	4.16(2.13)	4.44(2.77)	6.77(4.29)*					

Note: In order to improve the distributional characteristics of certain variables, the three TRI factors required square root transformations

* All MANOVA post-hoc comparisons (4) > (3,2,1) and (3,2) > (1) are statistically significant at $p < .01$.

Table 4

Descriptive statistics (means, standard deviations) of external validators and MANOVA post-hoc comparisons by DSM-V AUD diagnostic cutoffs

DSM-V AUD diagnosis	Overall	No AUD diagnosis(1)	AUD diagnosis(2)
	N = 396	N = 198(50%)	N = 198(50%)
External validator variables	Mean (SD)	Mean (SD)	Mean (SD)
Frequency of Alcohol Use Index	8.46(2.21)	7.67(2.21)	9.26(1.92)*
Quantity of Alcohol Use Index	7.81(2.24)	6.91(2.01)	8.69(2.11)*
CAPS-r sum scores	2.42(1.96)	1.52(1.49)	3.34(1.97)*
Govern (TRI)	6.57(4.52)	4.72(2.86)	8.43(5.07)*
Emotion (TRI)	7.06(4.25)	5.75(3.26)	8.37(4.71)*
Preoccupation (TRI)	4.48(2.85)	3.63(1.65)	5.31(3.47)*

Note: In order to improve the distributional characteristics of certain variables, the three TRI factors required square root transformations.

* All MANOVA post-hoc comparisons (2) > (1) are statistically significant at $p < .01$.

Table 5

Descriptive statistics (means, standard deviations) of external validators by DSM-V AUD severity qualifiers

DSM-V Qualifier Group	Overall	No/Minimal (1)	Moderate (2)	Severe (3)
	N = 396	N = 198 (50%)	N = 145 (36.16%)	N = 53 (13.38%)
External validator variables	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Frequency of Alcohol Use Index	8.46(2.21)	7.67(2.21)	9.01(1.81)	9.96(2.03) *
Quantity of Alcohol Use Index	7.81(2.24)	6.91(2.01)	8.32(1.84)	9.72(2.44) *
CAPS-r sum scores	2.42(1.96)	1.52(1.46)	2.84(1.79)	4.69(1.78) *
Govern (TRI)	6.57(4.52)	4.72(2.86)	7.42(4.24)	11.21(6.09) *
Emotion (TRI)	7.06(4.25)	5.75(3.26)	7.44(3.85)	10.91(5.84) *
Preoccupation (TRI)	4.48(2.85)	3.63(1.65)	4.79(2.99)	6.74(4.24) *

Note: In order to improve the distributional characteristics of certain variables, the three TRI factors required square root transformations.

* All MANOVA post-hoc group differences (3) > (2) > (1) are statistically significant at $p < 0.01$.