



Published in final edited form as:

*J Am Acad Child Adolesc Psychiatry*. 2008 November ; 47(11): 1329–1339. doi:10.1097/CHI.0b013e318184ff2e.

## Toward DSM-V: An Item Response Theory Analysis of the diagnostic process for DSM-IV Alcohol Abuse and Dependence in Adolescents

**Heather Gelhorn, Ph.D., Christie Hartman, Ph.D., Joseph Sakai, M.D., Michael Stallings, Ph.D., Susan Young, Ph.D., Soo Rhee, Ph.D., Robin Corley, Ph.D., John Hewitt, Ph.D., Christian Hopfer, M.D., and Thomas Crowley, M.D.**

Drs. Gelhorn, Hartman, Sakai, Hopfer and Crowley are with the School of Medicine at the University of Colorado Denver and Health Sciences Center; Drs. Stallings, Young, Rhee, Corley and Hewitt are at the Institute for Behavioral Genetics at the University of Colorado, Boulder.

Alcohol use, abuse, and dependence often first appear during adolescence<sup>1</sup>. The incidence of DSM-IV alcohol dependence begins near age 11, peaks at the age of 18, and declines rapidly from 18 to 25<sup>2</sup>. Adolescents and adults show different patterns of drinking behavior,<sup>3</sup> and may interpret the DSM-IV symptoms for alcohol abuse and dependence disorders (AUDs) differently and not necessarily the way these symptoms were intended<sup>4–6</sup>. Given these developmental trends in alcohol use, the differences between adults and adolescents, and the significant impact of these disorders it is important to study AUDs during adolescence.

A DSM-IV abuse diagnosis<sup>7</sup> requires endorsement of at least 1 of 4 abuse symptoms; a dependence diagnosis supersedes a diagnosis of abuse and requires endorsement of 3 or more of 7 dependence symptoms (clustered within a year). Although not specifically stated in the DSM, this diagnostic system might imply that abuse is less severe than dependence. The DSM-IV algorithm produces at least two problems. First, persons who endorse no abuse symptoms but 1 or 2 dependence symptoms do not qualify for a diagnosis, though they may have a greater severity of AUDs than those diagnosed with abuse. These persons have been termed “diagnostic orphans”<sup>8, 9</sup>. Second, persons who endorse all 4 abuse symptoms but only 2 dependence symptoms will be diagnosed with alcohol abuse, despite the fact that they have endorsed 6 total symptoms. In general, patterns of symptom endorsement that include 1 or 2 dependence symptoms and any number of abuse symptoms may lead to diagnoses that do little to distinguish severity of alcohol use disorder.

Item Response Theory (IRT)<sup>10, 11</sup> can be used to examine not only the abuse and dependence diagnoses, but the properties of the symptoms that comprise them. When examining a series of items such as the DSM-IV substance use disorder symptoms, rather than simply adding the number of symptoms endorsed by an individual, IRT examines which particular symptoms are endorsed by each individual, allowing that the symptoms may not have equal “weight” in predicting substance use problems. Thus, IRT estimates how informative a particular item is as well as how useful it is in predicting severity of pathology. Heuristically, it is helpful to think of IRT through the example of a test of mathematical ability. Instead of simply “counting” how many questions one answered correctly, IRT examines which questions were answered correctly and judges the “mathematical ability” of the test taker by this pattern. Similarly, each

Correspondence to Dr. Heather Gelhorn, 4200 East 9<sup>th</sup> Avenue, Box C268-35, Denver, CO, 80262; Heather.Gelhorn@UCHSC.edu. Article Plus online-only materials for this article appear on the Journal’s Web site: [www.jaacap.com](http://www.jaacap.com).

Disclosure: The authors report no conflicts of interest.

question would be rated as “harder” or “easier” based upon the patterns and frequency of correct answers. In the context of substance problems, certain symptoms may be endorsed only by more severely affected individuals who endorse many symptoms, whereas other symptoms may be endorsed by many, including those who endorse only one or two symptoms.

Research indicates that DSM-IV alcohol abuse and dependence are not diagnostically distinct entities, but represent a continuum of severity of AUDs<sup>12, 13</sup>, other studies have found that AUDs may represent two highly correlated factors<sup>14</sup>. Our study will examine alcohol abuse and dependence to answer two main questions important to an empirically-based revision of the DSM. First, do the current DSM-IV abuse and dependence diagnostic categories accurately and consistently distinguish between categories of AUD severity? If not, there may be alternative diagnostic algorithms that more effectively achieve this objective. Second, do the DSM-IV alcohol abuse and dependence symptoms characterize similar levels of AUD severity in clinical, community and adjudicated samples? The answer to this question will help to highlight symptoms which may require modification or reconsideration for DSM-V.

There have been several previous studies using IRT to examine alcohol use<sup>15, 16</sup>, and DSM abuse and dependence symptoms<sup>17–19</sup>. These previous studies have concluded that the severity parameters for the DSM-IV abuse and dependence symptoms are interspersed along the range of severity of AUD problems in both adults<sup>17, 19</sup> and adolescents<sup>18</sup>. An empirical test of the abuse-dependence distinction by Langenbucher et al. showed that the abuse and dependence item severity parameters could not be constrained to match the DSM-IV specification of abuse items as less severe than dependence items<sup>17</sup>.

Of particular relevance to the present study, is the study of DSM-IV alcohol symptoms in a clinical sample of 464 adolescents<sup>18</sup>. This study also concluded that the alcohol abuse and dependence items comprise a single unidimensional trait and that the item severity parameters were high for the symptoms *Withdrawal* and *Quit/Control*, but low for *Social Problems* and *Tolerance*. Items with better discrimination included *Neglect Roles*, *Time Spent*, and *Activities Given Up*, while *Hazardous Use* and *Quit/Control* had the poorest discrimination values. This study concluded that there is no clear distinction between abuse and dependence based on item severity parameters, and the authors suggest that further research be devoted to examining the utility and validity of a more dimensional approach to AUD diagnosis using abuse and dependence items together.

Previous studies have consistently concluded that the abuse-dependence distinction is not supported because the severity parameters of the abuse and dependence items are interspersed along the continuum of AUD severity. Despite this item severity parameter interspersion, it is still possible that the abuse and dependence constructs, as they are currently defined, might still adequately distinguish severity of AUD disorder based on the abuse and dependence diagnoses. The present study will extend the literature by examining how well the current DSM-IV algorithm and an alternative diagnostic algorithm distinguish AUD severity. The present study differs from previous studies in that it uses an extremely large community sample augmented with clinical cases. This combined clinical and community approach enriches the sample through adding subjects at the high-severity end of the continuum and yielding parameter estimates that are standardized to represent the distribution of AUDs in community adolescents. The present study aims to inform the design of future DSM diagnostic symptoms by providing information on the diagnostic categories and symptoms for AUDs during the critical period of adolescence.

## Method

### Participants

Participants included 5587 adolescents between the ages of 11–19 from three samples: clinical, adjudicated and community. Descriptive statistics are in Table 1. This sample is identical to that used to examine DSM-IV marijuana symptoms in a previous study<sup>20</sup>.

**Clinical sample**—Adolescents came from the family study component of the Center for Antisocial Drug Dependence (CADD), recruited between 1993–2003. Adolescents (n=615) between the ages of 12–18 were identified through university treatment centers in the Denver, Colorado area where they were receiving treatment for serious substance use and conduct problems. This sample has been described previously<sup>21</sup>.

**Adjudicated sample**—The Colorado Adjudicated Youth Study includes 179 youth aged 12–18 ascertained from state juvenile justice records and interviewed with the same assessment instruments as the other samples, and is described in more detail in<sup>22</sup>. The sample was collected between 2000 and 2007 and was 77% male with a mean age of 16.5 years (S.D. 1.5). Detailed information on a range of variables is available for the entire pool of potential subjects; the participants in our study are not significantly different from the larger pool of adjudicated youth on demographic, behavioral, and substance abuse variables<sup>22</sup>.

**Community sample**—The community sample includes 4793 participants between the ages of 12–18 from a variety of CADD projects including the Colorado Twin Registry, Colorado Adoption Project, and Family Control Samples. The twin sample consists of 3282 individuals from 1386 adolescent twin pairs and 510 siblings of twins recruited from the general population. The adoption sample consists of a total of 700 individuals including adoptees, matched controls, and their siblings. The Family Control sample includes 811 control probands and their siblings, the control probands were matched on age, sex, and ethnicity to adolescents from the clinical and adjudicated samples. Details of recruitment procedures and demographics are provided elsewhere<sup>23–25</sup>. For families with multiple participants the data were weighted for analyses so that each family comprised a single observation (e.g., in families with three siblings, the data were weighted .33, .33, .34). This weighting allows for extraction of the maximum amount of information by including the response patterns of all individuals while limiting each family in the analyses to a single observation.

### Measures and Procedures

Participants completed the Composite International Diagnostic Interview – Substance Abuse Module (CIDI-SAM), a valid and reliable structured interview which provides information on DSM-IV symptoms and diagnoses of abuse and dependence<sup>26–28</sup>. The present study uses lifetime symptoms and diagnoses based on information collected in the CIDI-SAM regarding the frequency, onset, and temporal clustering of DSM-IV substance use symptoms. Interviews were conducted by trained lay interviewers. Participants gave written informed assent/consent and were paid a nominal fee for participation. The research was approved by the Institutional Review Boards for each participating campus of the University of Colorado.

### Analyses

#### Factor Analyses

We tested for unidimensionality by conducting an exploratory factor analysis (EFA) of the alcohol symptoms using Mplus software<sup>29</sup>. In EFA, large ratios of first to second eigenvalues and a better fit of single versus multiple factor models are considered adequate evidence for unidimensionality<sup>17, 30</sup>.

## IRT analyses

The IRT model most appropriate for binary data such as substance abuse symptoms is the 2-parameter model<sup>10, 11</sup>. Under this model, 2 parameters are estimated for each symptom: severity (also known as “item difficulty” in the IRT literature) and discrimination. Symptom severity is defined as the severity point on the continuum of the latent AUD trait at which the probability of endorsing the symptom is 50%. Symptom severity tells us at what severity of disorder the symptom is most informative (i.e., is it most useful in scaling those with very serious or less serious AUDs?). Symptom discrimination represents the symptom’s ability to discriminate between persons with severity of disorder above and below the symptom’s threshold. The discrimination and severity parameters are similar to factor loadings and threshold parameters in a dichotomous factor analysis<sup>31</sup>.

IRT parameter estimation and Differential Item Functioning (DIF) analyses were conducted on the lifetime DSM-IV alcohol abuse and dependence symptoms using the software PARSCALE<sup>32</sup>. DIF tests examine whether each symptom indicates equivalent severity of disorder across clinical, adjudicated and community samples<sup>33</sup>. We did not conduct tests of DIF for the discrimination parameters due to power limitations.

## Evaluation of current DSM-IV diagnostic categories

IRT-based alcohol use severity scores can be estimated for each individual based on the number and the respective severity signals of endorsed symptoms. These IRT-based scores as well as total symptom counts were used to examine the DSM-IV diagnostic system. The severity of AUDs was compared across diagnostic orphans and persons with abuse or dependence diagnoses. Exploratory analyses were conducted to examine alternative means of categorizing AUDs. These exploratory analyses were aimed at eliminating the incongruities of diagnostic orphans and persons diagnosed with alcohol abuse who appear to have more severe AUDs.

## Results

### Prevalence Rates

The endorsement rates for each lifetime DSM-IV alcohol symptom are presented in Figure 1. Rates of alcohol abuse in the community, adjudicated and clinical samples were 10.0%, 23.8% and 41.3%, respectively. Alcohol dependence rates were 4.2%, 16.8% and 33.8%, respectively.

### Factor Analysis

Factor analysis of the DSM-IV abuse and dependence symptoms indicated that they comprise a single factor (RMSEA=0.06) that will be referred to as AUD severity. The first three eigenvalues were 4.81, 1.05, and 1.02 respectively, and the first factor accounted for 68.8% of the total variance. The factor structure was consistent across all samples and is consistent with previous reports of a single latent factor underlying AUDs<sup>34</sup>.

### Item Response Theory Analyses

The item characteristic curves (ICCs) for the lifetime DSM-IV alcohol abuse and dependence symptom parameters are shown in Figure 2. ICCs graphically display the severity and discrimination parameters and facilitate a visual comparison of parameters across symptoms. The x-axis for these figures represents the latent liability to AUDs expressed as z-scores standardized in the community sample. As shown in Figure 2, item discrimination parameters ranged from 0.61 (*Quit/Control*) to 1.51 (*Withdrawal*). Item severity parameters, which are scaled to represent z-scores standardized in the community sample, were lowest for *Tolerance* and highest for *Quit/Control*. The severity parameters were generally between 2 and 3.5 standard deviations above the mean of the community sample.

Figure 3 shows the results of the DIF analyses for the severity parameters. Two dependence (*Tolerance, Larger/Longer*) and one abuse item (*Neglect Roles*) displayed significant DIF across groups. After controlling for overall mean differences in severity, both *Tolerance* ( $\chi^2 = 19.68, p < .001$ ) and *Larger/Longer* ( $\chi^2 = 18.70, p < .001$ ) indicated significantly greater severity of AUDs in the clinical sample; the abuse item *Neglect Roles* indicated significantly lower severity of AUDs in the clinical sample ( $\chi^2 = 6.80, p = .03$ ). The DIF effect for this item was small and unlikely of great practical significance.

### Evaluation of current DSM-IV diagnostic categories

IRT severity and discrimination parameter estimates for each item were used to calculate AUD severity scores for each individual based upon their pattern of responses. Because the factor analysis showed that the DSM-IV abuse and dependence symptoms form a single factor, the abuse and dependence symptoms were used conjointly. IRT-based scores which consider the endorsed symptoms' severity signals are more precise than, but strongly correlated with, total symptom count. The increased precision results from considering the severity signal (see Figure 2) of the symptoms that were endorsed, in addition to the total number of symptoms. Table 2 shows average severity of AUDs displayed according to the DSM-IV diagnostic categories (i.e., no diagnosis, diagnostic orphans, abuse diagnosis, dependence diagnosis (with 12-month clustering)). IRT-based severity estimates indicate that those without a diagnosis are the least severe (mean = -0.22), diagnostic orphans are more severe (mean = 0.77), followed by abuse diagnoses (mean = 1.44) and dependence diagnoses (mean = 2.60). A test for mean severity differences across categories was highly significant (total symptom count:  $F=1729.22, p<.001, \eta^2=.70$ ; IRT-based severity estimates:  $F=1694.42, p<.001, \eta^2=.69$ ). These results suggest that, *on average*, the current diagnostic algorithm discriminates severity of AUDs.

However, subsequent analyses reveal that a substantial number of individuals are inaccurately categorized on the basis of AUD severity. The top two panels of Figure 4 display the distribution of individuals' AUD severity separated by diagnostic category for persons endorsing 1 or more abuse or dependence symptoms. The IRT-based severity estimates (top right panel) are scaled to have a mean of 0 and a standard deviation (SD) of 1 in the community sample and can be interpreted as standardized scores based on the community sample. The top panel of Figure 4 demonstrates that the current DSM-IV-based diagnostic system, though effective at distinguishing AUDs on average, results in substantial overlap and inconsistencies in severity across diagnostic categories. This problem is most clearly demonstrated in the top right panel of Figure 4 where IRT-based estimates of AUD severity for many of the diagnostic orphans are substantially higher (in some cases almost an entire SD higher) than severity estimates for many persons diagnosed with abuse. A similar pattern is evident across abuse and dependence. For example, some persons diagnosed with abuse have alcohol problem severity greater than +2 SDs, while many persons diagnosed with dependence have alcohol problem severity well below +2 SDs. A figure showing IRT-based estimates of alcohol use problem severity by total alcohol symptom count is available on the Journal's Web site at [www.jaacap.com](http://www.jaacap.com) through the Article Plus feature. With few exceptions, those who endorse 1 alcohol abuse or dependence symptom have AUD severity less than 1 SD greater than the mean. Persons who endorse 4 or more alcohol abuse or dependence symptoms largely have AUD severity greater than 2 SDs above the population mean.

### Exploratory Analyses – Alternative diagnostic algorithm

Subsequent exploratory analyses considered an alternative way to define diagnostic categories (see Table 2 and Figure 4 c and d). These analyses were guided by the results of the previous analyses suggesting that the alcohol abuse and dependence symptoms comprise a single factor, the severity signals of the symptoms are interspersed, and the current diagnostic system does not optimally reflect differences in AUD severity. Under the alternative diagnostic algorithm,

0 or 1 endorsed abuse or dependence symptoms results in no alcohol diagnosis, 2 or 3 symptoms results in an abuse diagnosis, and 4 or more symptoms results in a dependence diagnosis. Figures 4 c and d clearly indicate that all of the inconsistencies for total symptom count, and the vast majority of the inconsistencies in IRT-based severity of AUDs are eliminated under this alternative diagnostic algorithm. AUD severity is approximately 0–1 SDs above the mean for those without a diagnosis, 1–2 SDs above the mean for those with an abuse diagnosis, and 2 or more SDs above the mean for those with a dependence diagnosis.

The alternative algorithm results in improved discrimination across diagnostic categories as measured by both IRT-based severity estimates ( $F=4428.56$ ,  $p<.001$ ,  $\eta^2=.85$ ) and symptom count ( $F=2563.91$ ,  $p<.001$ ,  $\eta^2=.77$ ). As evidenced by comparing the  $\eta^2$  (i.e., explained variance) for the current DSM-IV algorithm versus the alternative algorithm, the difference in our ability to predict the severity of AUDs is marked, predicting 85% of the variance under the alternative diagnostic algorithm versus 69% under DSM-IV.

One concern with altering a diagnostic algorithm is the potential exists that some people who are diagnosed under the current system may no longer qualify for a diagnosis and therefore lose insurance coverage for their treatment. In our sample, the current DSM-IV diagnostic algorithm classifies 13.3% ( $n=742$ ) under alcohol abuse and 7.3% ( $n=408$ ) under dependence. The alternative algorithm classified 9.1% ( $n=506$ ) under abuse and 9.9% ( $n=554$ ) under dependence. Thus, while the proportion of alcohol problems remains relatively constant (20.6% ( $n=1150$ ) diagnosed under DSM-IV versus 19.0% ( $n=1060$ ) diagnosed under the alternative algorithm), as we show above, the severity of alcohol problems characterized by each diagnostic category is more consistent and informative under the alternative algorithm.

## Discussion

We conducted Item Response Theory-based analyses of DSM-IV lifetime alcohol abuse and dependence symptoms in 5587 community, clinical, and adjudicated adolescents. The results suggest that an alternative diagnostic algorithm may provide a clearer severity-based distinction between diagnostic categories. The results of this study may be informative in the development of the upcoming DSM-V.

### DSM-IV alcohol abuse and dependence item characteristics

IRT analyses can characterize the severity signal of individual DSM-IV symptoms. Our results are highly consistent with those of a previous study of DSM-IV alcohol symptoms in adolescents with severity parameters that were low for *Tolerance* and *Social Problems*, but high for *Withdrawal* and *Quit/Control*<sup>18</sup>. Also consistent with previous reports, the symptoms *Neglect Roles*, *Time Spent*, and *Activities Given Up* had higher discrimination values, while *Quit/Control* and *Withdrawal* were less discriminating<sup>4, 18</sup>. *Quit/Control* has been identified as problematic because many teens do not meet symptoms for dependence until they make a positive behavioral change by attempting to quit or control their use<sup>35</sup>. Notably, our results were highly consistent with the previous study of adolescents though each study used a different structured clinical interview to assess DSM-IV AUD symptoms<sup>18</sup>. As displayed in Figure 2, the alcohol abuse and dependence symptoms are most informative at levels of severity greater than 2 SDs above the mean, suggesting that the symptoms are most informative in adolescents with very severe alcohol problems as compared with what is observed in the general population (i.e., standardized to the community sample).

### Abuse-dependence Distinction

The DSM-IV distinguishes between alcohol abuse and dependence, suggesting that abuse indicates more recent initiation of the substance. The abuse diagnosis is preempted by a

dependence diagnosis, suggesting that the latter is more severe<sup>7</sup>. We examined this categorical distinction between abuse and dependence based on severity. The abuse-dependence distinction is not supported by our results at the symptom level (Figure 2) nor at the diagnostic level (Figure 4). At the symptom level, the severity parameters for the abuse symptoms are interspersed among the dependence symptoms and not clustered along the less severe end of the spectrum, consistent with previous studies<sup>17–19</sup>. At the diagnostic level, adolescents categorized into DSM-IV abuse versus dependence differ significantly in their IRT-based estimates of AUD severity. However, there are considerable inconsistencies in the AUD severity of many individuals across DSM-IV diagnostic categories. Taken together, these data suggest that the abuse-dependence distinction currently outlined in DSM-IV might be reconsidered for DSM-V. We provide evidence that a diagnostic system that considers all the abuse and dependence items together will more accurately scale adolescents on AUD severity. The proposed alternative diagnostic algorithm results in more accurate prediction of severity of AUDs across diagnostic category (based on explained variance), and lower variability in AUD severity within diagnostic categories (bottom panels Figure 4) compared with the current DSM-IV diagnostic algorithm. The alternative algorithm does not change the severity of abuse and dependence diagnoses on average (Table 2). In addition to the benefit of more precise measurement of AUD severity, the alternative algorithm is advantageous because it eliminates diagnostic orphans, a frequently cited problem with the current abuse-dependence distinction<sup>8, 9, 36, 37</sup>. Adolescent diagnostic orphans are troubling because they have high rates of other drug use and clinically significant problems commensurate with their counterparts who qualify for diagnoses<sup>38</sup>. Additionally, defining abuse and dependence as categorical classifications derived from a continuous measurement scale could be beneficial for studies that require reliable and accurate characterization of the severity of alcohol use problems. Some examples include genetic studies where power is greater with continuous phenotypes, and studies on the development of alcohol dependence.

## DIF Results

The results of our DIF analyses (Figure 3) suggest that two dependence symptoms (*Tolerance* and *Larger/Longer*) display differences in severity between samples. These results are consistent with at least two alternative hypotheses. First, persons who are predisposed to alcohol problems may have a lower initial level of sensitivity and response to alcohol<sup>39</sup>; thus, persons with severe problems may actually be experiencing (and thus reporting) these symptoms less than others. Alternatively, *Tolerance* and *Larger/Longer* may be more common in adolescents with limited drinking experience because they are developing pharmacological tolerance as their drinking progresses<sup>4, 6</sup>. Teens do not typically report any desire to limit the amount or length of their alcohol consumption, and thus, endorse *Larger* only because they drank more than on previous occasions<sup>4</sup>. In a sample of adults, *Larger/Longer* was the only dependence item that did not significantly relate to problem outcomes at 5-year follow-up<sup>40</sup>. Consistent with prior evidence of the unreliability of this symptom, our results show that *Larger/Longer* is of much lower severity in adjudicated adolescents and also significantly different across community and clinical samples. An IRT-based study of DSM-IV symptoms for cannabis use disorders in this same sample also identified DIF in these two dependence items<sup>41</sup>. A future IRT-based study of clinical and community adults might resolve whether these symptom differences are specific to adolescence. These items may warrant reconsideration in DSM-V because it is undesirable to have diagnostic symptoms that are less likely to be endorsed by clinically relevant persons, or are consistently misinterpreted by adolescents who are at greatest risk of developing these disorders.

## Implications for DSM-V

- 1. Alcohol abuse and dependence symptoms should be considered conjointly**  
—Taken together with previous factor analyses supporting a common latent factor underlying

alcohol abuse and dependence, this study suggests that DSM-V should consider abuse and dependence symptoms conjointly with separate thresholds for abuse and dependence diagnoses. This study outlines one potential alternative algorithm. Although this represents a departure from the theoretical framework underlying these two diagnoses<sup>42</sup>, such an approach would have the appeal of eliminating diagnostic orphans, and our results show that this approach is empirically supported by improved scaling of AUD severity within and across diagnostic categories.

**2. Tolerance and Larger/Longer do not have consistent severity signals across samples**—These findings should be studied further. If this phenomenon represents a biological difference<sup>39</sup> and those at greatest risk for serious alcohol dependence are less likely to develop these symptoms, then they should be reconsidered. If this phenomenon reflects differential interpretation of the questions, then the questions may need rewording.

The results of the present study should be considered in view of the following limitations. First, this sample consisted of adolescents from Colorado, USA. This is not an epidemiological sample, and it is unclear to what extent the results can be generalized to adults or to other geographic areas. The symptoms and diagnoses used in the analyses were based on the self-reports of adolescents participating in a direct interview. Our adjudicated and clinical samples included mostly male subjects, while our community sample was mostly Caucasian. However, our item-level results were largely consistent with results from other studies of alcohol abuse and dependence symptoms<sup>17, 19</sup>. It is important to note that treated cases were not excluded from the community sample, and therefore reflect their incidence in the underlying population. Due to power limitations, we were unable to test for differences related to sex, ethnicity or other variables. The DSM-IV alcohol symptoms were assessed based on self-reports and one operationalization of the DSM-IV symptoms. Our adjudicated sample included only 179 subjects; although this is a respectable size, the estimates for this group should be interpreted with caution. The abuse and dependence thresholds that we propose are based on severity of AUDs and the utility of these diagnostic categories for prognostic or treatment response information is unknown. Finally, our data did not include an alternative measure of AUD severity that might help assess the alternative algorithm without using information from the current symptoms. In future studies, we hope to examine follow-up data on the current sample.

## Acknowledgments

Financial support was provided through grants DA11015, DA12845, DA05131, DA015522, MH01865, DA016314

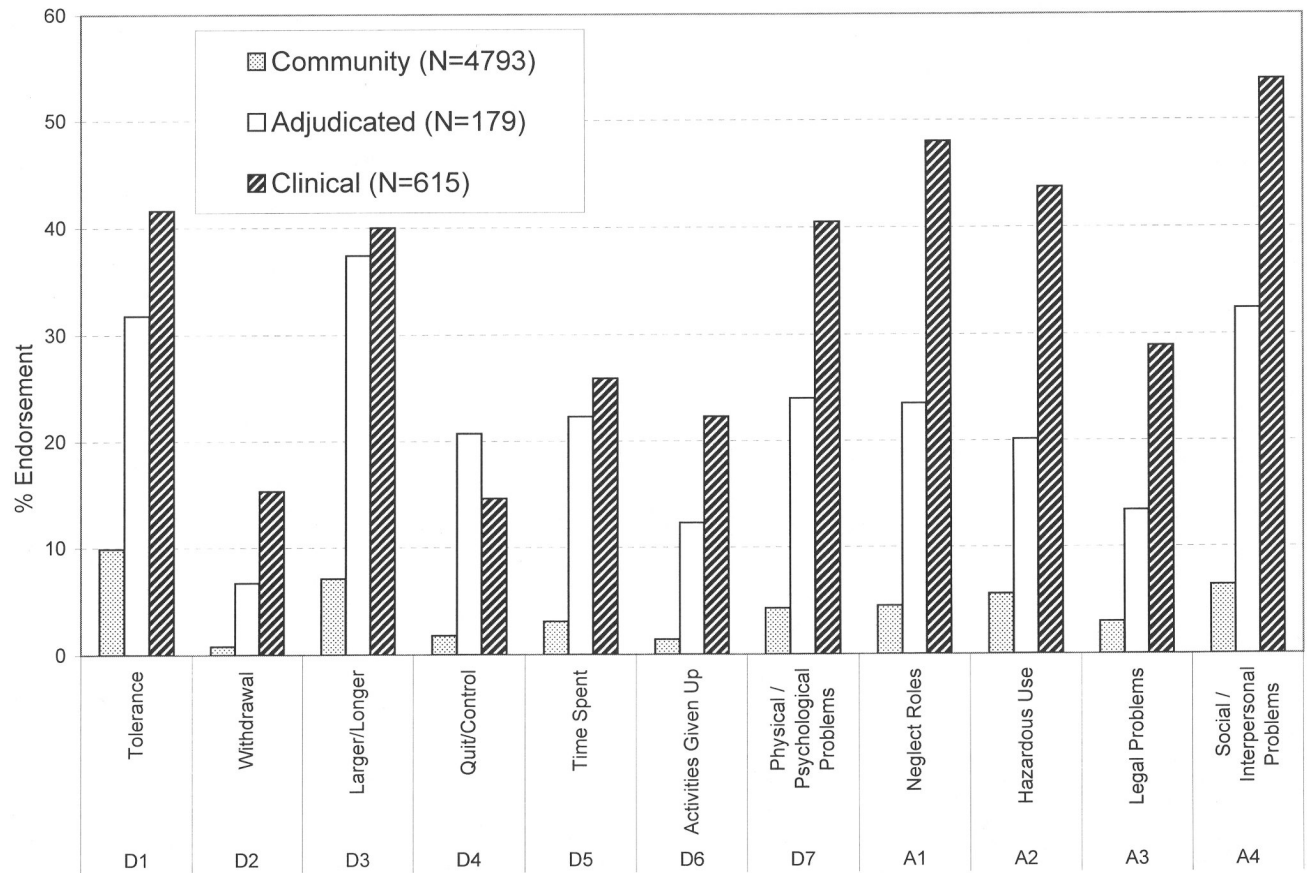
## References

1. 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(3):309–322. [PubMed: 12393225]
2. Li TK, Hewitt BG, Grant BF. Alcohol use disorders and mood disorders: a National Institute on Alcohol Abuse and Alcoholism perspective. *Biol Psychiatry* 2004 Nov 15;56(10):718–720. [PubMed: 15556112]
3. Martin CS, Kaczynski NA, Maisto SA, Bukstein OM, Moss HB. Patterns of DSM-IV alcohol abuse and dependence symptoms in adolescent drinkers. *J Stud Alcohol* 1995 Nov;56(6):672–680. [PubMed: 8558899]
4. 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 Nov 1;80(2):191–200. [PubMed: 15894432]
5. Chung T, Martin CS, Winters KC. Diagnosis, course, and assessment of alcohol abuse and dependence in adolescents. *Recent Dev Alcohol* 2005;17:5–27. [PubMed: 15789857]

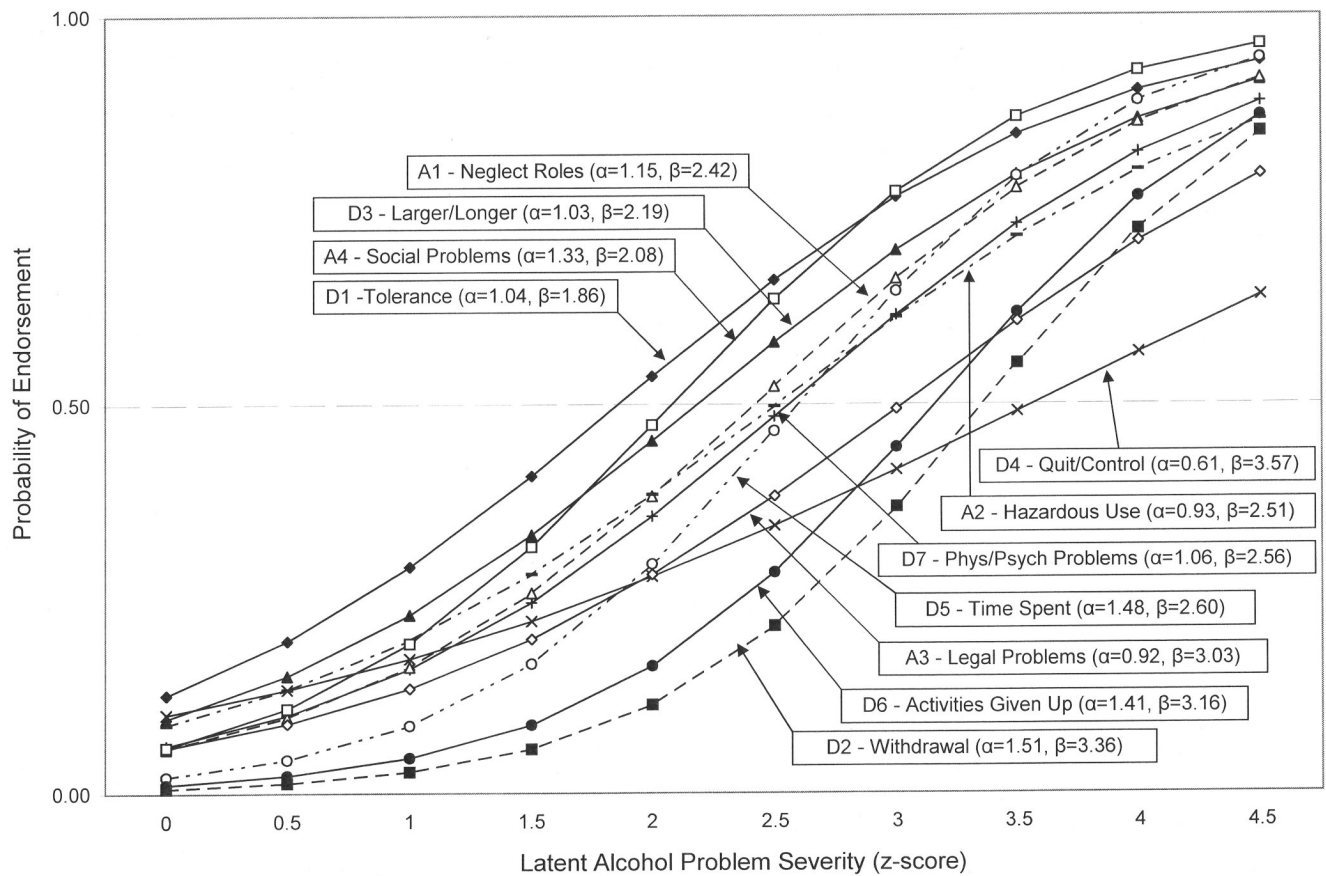


6. Chung T, Martin CS, Winters KC, Langenbucher JW. Assessment of alcohol tolerance in adolescents. *J Stud Alcohol* 2001 Sep;62(5):687–695. [PubMed: 11702808]
7. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Vol. 4th Edition (DSM-IV). Washington, DC: American Psychiatric Association; 2004.
8. Harrison PA, Fulkerson JA, Beebe TJ. DSM-IV substance use disorder criteria for adolescents: a critical examination based on a statewide school survey. *Am J Psychiatry* 1998 Apr;155(4):486–492. [PubMed: 9545993]
9. Hasin D, Paykin A. Dependence symptoms but no diagnosis: diagnostic 'orphans' in a community sample. *Drug Alcohol Depend* 1998 Mar 1;50(1):19–26. [PubMed: 9589269]
10. Lord, FM.; Novick, MR. *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley; 1968.
11. Embretson, SE.; Reise, SP. *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 2000.
12. Li TK, Hewitt BG, Grant BF. The Alcohol Dependence Syndrome, 30 years later: a commentary. the 2006 H. David Archibald lecture. *Addiction* 2007 Oct;102(10):1522–1530. [PubMed: 17680851]
13. Bucholz KK, Heath AC, Reich T, et al. 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 Nov;20(8):1462–1471. [PubMed: 8947326]
14. Muthen BO. Factor analysis of alcohol abuse and dependence symptom items in the 1988 National Health Interview Survey. *Addiction* 1995 May;90(5):637–645. [PubMed: 7795500]
15. Kahler CW, Strong DR, Stuart GL, Moore TM, Ramsey SE. Item functioning of the alcohol dependence scale in a high-risk sample. *Drug Alcohol Depend* 2003;72:183–192. [PubMed: 14636973]
16. Krueger RF, Nichol PE, Hicks BM, et al. Using latent trait modeling to conceptualize an alcohol problems continuum. *Psychol Assess* 2004 Jun;16(2):107–119. [PubMed: 15222807]
17. Langenbucher JW, Labouvie E, Martin CS, et al. An application of item response theory analysis to alcohol, cannabis, and cocaine criteria in DSM-IV. *J Abnorm Psychol* 2004 Feb;113(1):72–80. [PubMed: 14992659]
18. 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 Nov;115(4):807–814. [PubMed: 17100538]
19. 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 Jul;36(7):931–941. [PubMed: 16563205]
20. Hartman, C.; Gelhorn, HL.; Sakai, JT., et al. Paper presented at: College on Problems of Drug Dependence. Phoenix, AZ: 2006. An item response theory analysis of DSM-IV marijuana abuse and dependence criteria in adolescents.
21. Stallings MC, Corley RP, Hewitt JK, et al. A genome-wide search for quantitative trait loci influencing substance dependence vulnerability in adolescence. *Drug Alcohol Depend* 2003;70:295–307. 2003. [PubMed: 12757967]
22. Hopfer, CJ.; Salomonsen-Sautel, S.; Corley, RP., et al. Drug dependence among adjudicated youth. College of Problems of Drug Dependence. 2007 [Accessed May 2, 2008]. <http://www.cpdd.vcu.edu>
23. Plomin, R.; Campos, J.; Corely, RP., et al. Individual differences during the second year of life: the MacArthur Longitudinal Twins Study. In: Colombo, J.; Fagan, J., editors. *Individual differences in infancy: reliability, stability, prediction*. Hillsdale, NJ: Lawrence Erlbaum; 1990. p. 431-455.
24. Rhea SA, Gross AA, Haberstick BC, Corley RP. Colorado Twin Registry. *Twin Res Hum Genet* 2006 Dec;9(6):941–949. [PubMed: 17254434]
25. Miles DR, Stallings MC, Young SE, Hewitt JK, Crowley TJ, Fulker DW. A family history and direct interview study of the familial aggregation of substance abuse: the adolescent substance abuse study. *Drug Alcohol Depend* 1998;49(2):105–114. [PubMed: 9543647]
26. Crowley TJ, Mikulich SK, Ehlers KM, Whitmore EA, MacDonald MJ. Validity of structured clinical evaluations in adolescents with conduct and substance problems. *J Am Acad Child Adolesc Psychiatry* 2001;40(3):265–273. [PubMed: 11288767]

27. Cottler LB, Robins LN, Helzer JE. The reliability of the CIDI-SAM: a comprehensive substance abuse interview. *Br J Addict* 1989 Jul;84(7):801–814. [PubMed: 2758153]
28. Robins LN, Wing J, Wittchen HU, et al. The Composite International Diagnostic Interview. An epidemiologic Instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry* 1988 Dec;45(12):1069–1077. [PubMed: 2848472]
29. Mplus [computer program]. Version. Los Angeles, CA: Muthen & Muthen; 1998.
30. Krueger RF, Finger MS. Using item response theory to understand comorbidity among anxiety and unipolar mood disorders. *Psychol Assess* 2001 Mar;13(1):140–151. [PubMed: 11281035]
31. Takane Y, de Leeuw J. On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika* 1987;52(3):393–408.
32. PARSCALE [computer program]. Version. Chicago: 2003.
33. Thissen D, Steinberg L, Gerrard M. Beyond group mean differences: The concept of item bias. *Psychol Bull* 1986;99:118–128.
34. Proudfoot H, Baillie AJ, Teesson M. The structure of alcohol dependence in the community. *Drug Alcohol Depend* 2006 Jan 4;81(1):21–26. [PubMed: 16005578]
35. Martin CS, Fillmore MT, Chung T, Easdon CM, Miczek KA. Multidisciplinary perspectives on impaired control over substance use. *Alcohol Clin Exp Res* 2006 Feb;30(2):265–271. [PubMed: 16441275]
36. Lewinsohn PM, Rohde P, Seeley JR. Alcohol consumption in high school adolescents: frequency of use and dimensional structure of associated problems. *Addiction* 1996 Mar;91(3):375–390. [PubMed: 8867200]
37. Hasin D, Paykin A. Dependence symptoms but no diagnosis: diagnostic 'orphans' in a 1992 national sample. *Drug Alcohol Depend* 1999 Feb 1;53(3):215–222. [PubMed: 10080047]
38. Pollock NK, Martin CS. Diagnostic orphans: adolescents with alcohol symptom who do not qualify for DSM-IV abuse or dependence diagnoses. *Am J Psychiatry* 1999 Jun;156(6):897–901. [PubMed: 10360129]
39. Schuckit MA. Low level of response to alcohol as a predictor of future alcoholism. *Am J Psychiatry* 1994;151(2):184–189. [PubMed: 8296886]
40. Schuckit MA, Danko GP, Smith TL, Buckman KR. The five-year predictive validity of each of the seven DSM-IV items for alcohol dependence among alcoholics. *Alcohol Clin Exp Res* 2002 Jul;26(7):980–987. [PubMed: 12170107]
41. Hartman CA, Gelhorn H, Crowley TJ, et al. Item response theory analysis of DSM-IV cannabis abuse and dependence criteria in adolescents. *J Am Acad Child Adolesc Psychiatry* 2008 Feb;47(2):165–173. [PubMed: 18176333]
42. Edwards G, Gross MM. Alcohol dependence: provisional description of a clinical syndrome. *Br Med J* 1976 May 1;1(6017):1058–1061. [PubMed: 773501]

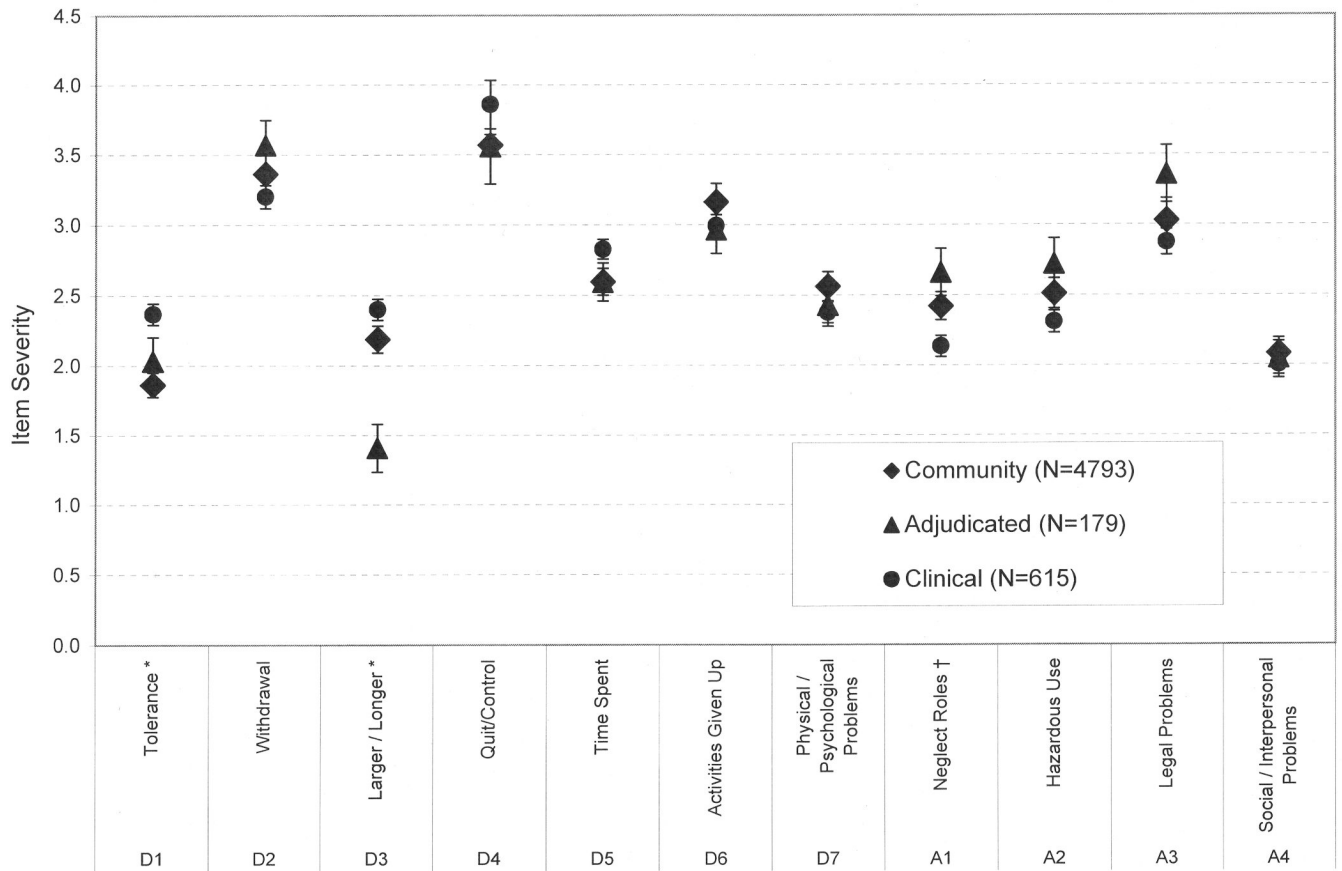


**Figure 1. Percent endorsement of DSM-IV lifetime abuse and dependence symptoms in three groups**  
 D1-D7 and A1-A4 are the symptoms of dependence and abuse (respectively) as listed in DSM-IV.



**Figure 2. Item characteristic curves for each of the DSM-IV lifetime alcohol abuse and dependence symptoms**

$\alpha$  = item discrimination parameter, analogous to a factor loading, the slope of the line at the point where the probability of endorsing the item is 50%;  $\beta$  = item severity parameter, the point on the standardized latent alcohol problem severity scale where the probability of endorsing the symptom is 50%; items that are further to the right of the diagram indicate more severe pathology. D1–D7 and A1–A4 are the alcohol dependence and abuse symptoms (respectively) from DSM-IV. The x-axis represents z-scores for AUD problems standardized to the community sample.

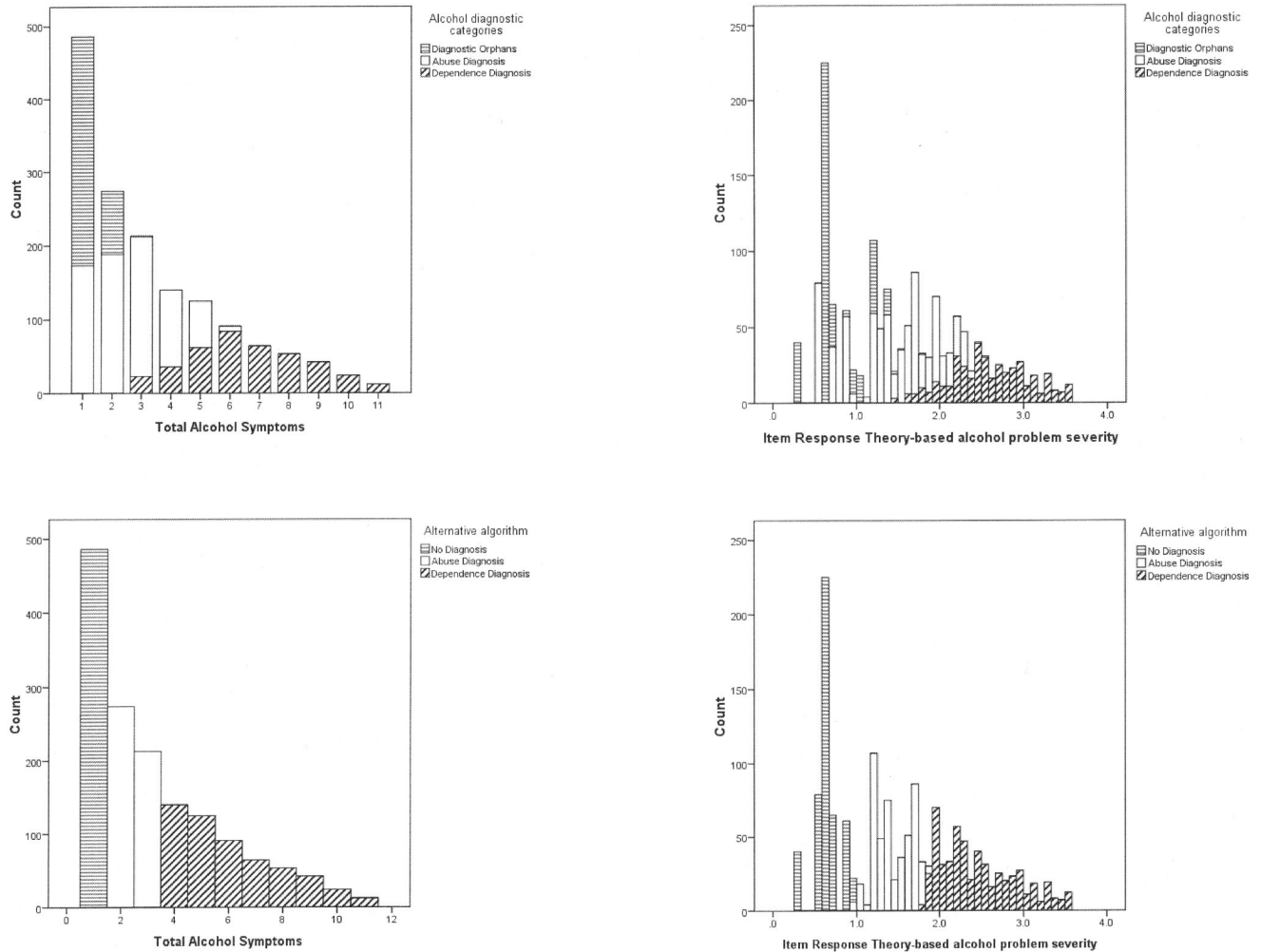


**Figure 3. Differential Item Functioning (DIF) results - Item severity parameters and the standard error for each DSM-IV lifetime alcohol abuse and dependence symptom**

Item severity parameters are displayed on a scale of alcohol problem severity that has a mean of 0 and standard deviation of 1 in the general population

\* item displayed significant Differential Item Functioning (DIF) with  $p < .001$

‡ item displayed significant DIF with  $p < .05$



**Figure 4. Proportion of individuals in each DSM-IV diagnostic category and their corresponding alcohol problem severity**

Figure 4 displays the observed frequency for (a) each person’s total alcohol symptom count (top left panel) and (b) IRT-based estimate of alcohol use problem severity (top right panel) for persons in the study who endorsed one or more symptoms. For clarity, the 4063 persons endorsing no DSM-IV abuse or dependence symptoms are not represented in the figures. The shading indicates the proportion of individuals who are: horizontal stripes - diagnostic orphans (endorsing 1 or 2 dependence symptoms but no abuse symptoms), no shading - diagnosed with abuse, and diagonal stripes - diagnosed with dependence (with clustering criterion) under the current DSM-IV diagnostic algorithm. The bottom panel displays the results under the alternative algorithm (1 symptom = no diagnosis, 2 or 3 symptoms = abuse diagnosis, 4 or more symptoms = dependence diagnosis) for (c) total alcohol symptom count (bottom left panel) and (d) IRT-based estimate of alcohol use problem severity (bottom right panel) for persons in the study who endorsed one or more symptoms. The shading indicates the proportion of individuals who do not qualify for a diagnosis (horizontal stripes), diagnosed with abuse (no shading), and diagnosed with dependence (diagonal stripes; with clustering criterion) under the proposed alternative diagnostic algorithm.

Table 1

## Sample demographics

Sample	N	Mean Age	Gender	Ethnicity
Combined	5587	15.3 (2.1)	M = 58% F = 42%	Caucasian: 78% Hispanic: 14% African American: 3% Other: 5%
Clinical	615	15.8 (1.2)	M = 92% F = 8%	Caucasian: 50% Hispanic: 37% African American: 9% Other: 4%
Adjudicated	179	16.3 (1.6)	M = 77% F = 23%	Caucasian: 63% Hispanic: 14% African American: 10% Other: 13%
Community	4793	15.2 (2.2)	M = 53% F = 47%	Caucasian: 82% Hispanic: 11% African American: 2% Other: 5%

Note. Standard deviation in parentheses; Other = other ethnicity or biracial.

Table 2

Severity of alcohol use problems under DSM-IV and alternative diagnostic algorithms

DSM-IV Diagnostic Algorithm (with clustering for dependence)							
Diagnostic Category	n	Abuse Symptoms Met	Dependence Symptoms Met	Mean No. Ab and Dep Symptoms Met	Sum of Endorsed Items' Severity Parameters (a)	Mean of Endorsed Items' Severity Parameters (b)	IRT-based Alcohol Problem Severity
No Diagnosis	4463	0	0 to 2	0.11 (0.371)	0.25 (0.88)	0.21 (0.68)	-0.22 (0.32)
Diagnostic Orphans (c)	400	0	1 or 2	1.22 (0.421)	2.82 (1.17)	2.30 (0.53)	0.77 (0.29)
Abuse	724	1 to 4	Less than 3	2.61 (1.282)	6.23 (3.15)	2.38 (0.23)	1.44 (0.51)
Dependence	400	0 to 4	At least 3	6.62 (2.00)	16.60 (5.52)	2.49 (0.14)	2.60 (0.48)

Alternative Diagnostic Algorithm							
Diagnostic Category	n	Abuse and Dependence Symptoms Met	Mean No. Ab and Dep Symptoms Met	Sum of Endorsed Items' Severity Parameters (a)	Mean of Endorsed Items' Severity Parameters (b)	IRT-based Alcohol Problem Severity	
No Diagnosis	4549	0 or 1	0.11 (0.31)	0.25 (0.74)	0.25 (0.74)	-0.21 (0.30)	
Abuse	487	2 or 3	2.44 (0.50)	5.76 (1.34)	2.36 (0.26)	1.44 (0.22)	
Dependence	551	At least 4	6.09 (1.92)	15.14 (5.29)	2.47 (0.14)	2.47 (0.46)	

(a) Total severity scores of all symptoms present, averaged across subjects

(b) Mean severity scores of all symptoms present, average across subjects

(c) Diagnostic Orphans comprise a subcategory of "No Diagnosis"