



Published in final edited form as:

Am J Addict. 2013 July ; 22(4): 402–410. doi:10.1111/j.1521-0391.2013.12042.x.

Alcohol Use Disorders among Substance Dependent Women on Temporary Assistance with Needy Families: More Information for Diagnostic Modifications for DSM-5

Alexis N. Kuerbis, LCSW, PhD^{1,2}, Brett T. Hagman, PhD³, and Jon Morgenstern, PhD^{1,2}

¹Research Foundation for Mental Hygiene, Inc., New York, New York

²Columbia University College of Physicians and Surgeons, New York, New York

³University of South Florida, Tampa, Florida

Abstract

Purpose—While modifications to alcohol use disorder (AUD) criteria are proposed for DSM-5, examination of the criteria’s performance among highly vulnerable populations is lacking. This study determined the dimensionality and rank order severity of the DSM-IV AUD criteria among Temporary Assistance for Needy Families (TANF) recipients with high rates of chemical dependency and co-morbid mental health disorders.

Method—Secondary analysis was performed on data from 461 TANF eligible women screened for AUD criteria using the Structured Clinical Interview for DSM-IV-TR. Exploratory (EFA) and confirmatory factor analyses (CFA) were performed on the AUD criteria. Two-parameter Item Response Theory (IRT) analysis was performed to determine item location and discrimination of criteria for both abuse and dependence. Differential Item Functioning for those with an additional substance use disorder or with high levels of depressive symptoms was explored.

Results—41.2% met criteria for dependence, and 4.4% for abuse. EFA and CFA revealed a two-factor model provided adequate fit to criteria, and IRT indicated a potential hierarchical order between the criteria--abuse being more severe but dependence having greater reliability.

Conclusion—Contrary to existing literature, findings suggest that a two factor solution may more be more appropriate. Implications are discussed.

As of the third edition of the *Diagnostic and Statistical Manual of Mental Disorders*, alcohol use disorder (AUD) criteria were conceptualized on the basis of the Alcohol Dependence Syndrome¹, a dimensional construct of drinking impairment^{2,3}. Rather than represent AUDs as a continuous construct, they were split into two primary diagnoses, alcohol abuse (AA) and alcohol dependence (AD). These two factors were thought to represent a hierarchical relationship in which AD was: (a) more severe than AA³ and (b) distinguished from AA criteria by physical dependence symptoms². Furthermore, neither diagnosis has an indicator of severity, creating an “artificial threshold” (i.e., either presence or absence of the disorder)³ in research and clinical settings alike in which potential important information about an individual’s particular AUD presentation, and severity, is lost.

Address correspondence to Dr. Kuerbis, c/o Columbia Addiction Services and Psychotherapy Research, 3 Columbus Circle, Suite 1404, New York, NY 10019. Kuerbis@nyspi.columbia.edu.

Declaration of Interest:

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this paper.

While this hierarchical relationship was maintained through the current *Diagnostic and Statistical Manual of Mental Disorders IV*, text revision (DSM-IV)⁴, there is mounting evidence that modifications to the criteria and their presumed hierarchical relationship are required for DSM-5. Results of some studies demonstrated that AUDs are best represented by two distinct factors⁵⁻⁷; however, the majority of these studies revealed a valid and strongly supported AD factor, whereas AA remained equivocal with mixed results. Several other psychometric examinations of the criteria including exploratory and confirmatory factor analysis⁸⁻¹¹, latent class analysis¹², and Rasch and item response theory analysis¹³ have demonstrated that abuse and dependence diagnoses together may actually describe the same disorder along a continuum of severity. Furthermore, evidence of AD as a more severe disorder or AA as prodromal to dependence is lacking. As a result of the above findings, proposed modifications to the AUD criteria for DSM-5¹⁴ involve an integration of the two disorders into a single disorder, representing a single construct of AUD. Additionally, symptom counts, as in counting the number of criteria satisfied, are proposed to determine severity specifiers.

Research investigating the validity of combining these two disorders into a single diagnosis for AUD has a key limitation. There has been little to no research on highly complex clinical populations with AUDs and how modifying diagnostic criteria might alter their identification, and classification. Research has largely examined the dimensionality of these criteria in general adult populations^{10,11,13,15}, late adolescents^{8,16} and college students^{17,18}. Among the very few studies to examine an adult clinical population, e.g.,^{15,19}, relatively few women and ethnic minorities were included in the samples. As such, a precise examination of how these criteria manifest in frequency and severity across specific, highly vulnerable populations is imperative.

Women on Temporary Assistance to Needy Families (TANF) are a unique population with significantly higher rates of co-morbid substance use disorders (SUDs) and mental health disorders than the general population, and they represent some of the most complex cases of chemical dependency²⁰. Women on TANF have close to twice the rates of multiple substance dependence²¹ and co-morbid mental health disorders than in the general population or their non-recipient counterparts^{22,23}; the most common of which are major depressive disorder and dysthymia^{24,25}. Existence of co-occurring disorders may impact the dimensionality of AUD in previously undetermined ways. Furthermore, women on welfare are significantly less likely to receive specialty mental health or substance abuse treatment than the general population²³, and understanding how we identify and classify such disorders are one step toward addressing this unmet need. Screening and diagnoses are the initial tools used in welfare offices to refer and engage TANF women in treatment²⁶⁻²⁸, and thus research on TANF recipients provides an important opportunity to explore the performance of the diagnostic criteria of AUDs in the context of this clinical and psychosocial complexity.

This study aimed to investigate the dimensionality and the internal structure and rank order of severity of the DSM-IV AUD criteria in a population of TANF recipients using exploratory and confirmatory factor analysis and Item Response Theory (IRT) methods on participants' responses to the Structured Clinical Interview for DSM-IV (SCID)²⁹. That is, we investigated whether TANF recipients endorsed items that represent diagnostic criteria of AA and AD in such a way that suggest the existence of one underlying construct—specifically, AUD--on a continuum of severity. Additionally, we explored how each of the criteria of AA and AD performed in relation to one another in terms of denoting severity of AUD among TANF recipients. For example, for more a more vulnerable population, are AD criteria more severe than AA criteria? In order to explore these questions, we conducted a secondary analysis on screening data collected from the Substance Abuse Research

Demonstration (SARD) project^{30,31}, a welfare demonstration project investigating the effectiveness of intensive case management in facilitating receipt of substance abuse treatment, employment and exit from welfare for TANF recipients.

METHOD

Subjects

Five hundred and ninety-five women screened positive for at least some substance use and were potentially eligible for the SARD study. Of those, 461 TANF eligible women completed the SCID, the primary measure used in this analysis. There is a subset of women (n=159) who are included in the present analysis but not included in the final SARD sample (N=302) due to a variety of reasons including refusing to participate, not qualifying for alcohol or other drug dependence, seeking residential or already receiving substance abuse treatment. Demographic characteristics of this sample did not differ from the final SARD sample (n=302), the details of which are reported elsewhere^{30,31} and reviewed here briefly. Participants were between the ages of 18 and 54 years; 92.3% were non-Hispanic Black, 4.5% were Hispanic, and 3.2% were other--a representative sample of Essex County welfare recipients who satisfied eligibility criteria³². The typical participant was around 34 years old, had a mean of nine years on welfare, had between two and three children, and had never been married.

Within the present sample (N=461), all women endorsed at least one substance abuse or dependence item. Under half (42.1%) met criteria for AD; 45.3% cocaine dependence; 33.2% opioid dependence; and 9.8% cannabis dependence. Sixty-six percent of those with AD also qualified for one of the above listed SUDs. Thus this sample was comprised of a large proportion of participants with multiple substance dependence. Specifically, among those with AD, 51% met criteria for cocaine dependence, 26.8% met criteria for opioid dependence, and 8.8% met criteria for cannabis dependence. Under five percent (4.4%) met criteria for AA, all of whom (except for two individuals) had a primary other drug dependence diagnosis. While all participants endorsed at least one criterion for AD or AA, 16% did not satisfy criteria for any alcohol or other drug use disorder. For example, among these diagnostic orphans, 5.6% of participants met one to two criteria for AD and did not qualify for AA.

Recruitment and Screening

Individuals were approached in welfare offices in Essex and Atlantic County, NJ while applying for or seeking recertification of TANF benefits. Once participants consented to participation, they were referred for a full diagnostic evaluation with an addictions counselor. All interviews took place in a location that protected participant privacy and right to confidentiality.

Eligibility criteria—Formal selection criteria for participants in SARD included being (1) TANF eligible, (2) eligible for New Jersey's welfare-to-work program, and (3) able to speak English well enough to complete an interview³². Women were excluded from this particular analysis if they were: (1) actively psychotic or receiving treatment for a psychotic disorder or (2) medically deferred from work requirements.

Measures

Demographic information, such as age, race, ethnicity, and marital status were collected using the Addiction Severity Index–Expanded Female Version³³. This instrument is the standard of structured interviews for evaluating individuals with SUDs in both research and clinical settings and demonstrates high reliability and validity across numerous studies³⁴.

The Structured Clinical Interview for DSM-IV (SCID)²⁹ is a structured interview that was implemented by addiction counselors as part of an in-depth evaluation of substance use diagnoses. Specifically, the SCID was used to assess for endorsement of alcohol and other drug abuse and dependence criteria during the initial interview that screened for study eligibility. For the current study, the 11 AA and AD criteria were the focus of these analyses. Additionally, a dichotomous variable was created that indicated whether a participant met criteria for an additional SUD (e.g., cannabis, cocaine, and opioid).

Depressive symptoms were measured using the Beck Depression Inventory, Second Edition (BDI-II)³⁵. The BDI-II is a self-report, 21-item questionnaire, which yields a continuous score, ranging from 0 to 63 indicating level of experienced distress associated with depressive symptoms³⁵. The reliability and validity of the BDI-II with populations with SUDs is well established (e.g.,^{36,37}), with a high degree of internal consistency in this sample (i.e., Cronbach's Coefficient alpha = .908). Previous studies demonstrate that scores of 20 and above on the BDI-II are sensitive and specific in indicating the presence of major depressive disorder³⁵, and 20 is a widely-used threshold to indicate presence of clinically significant depressive symptoms across research studies³⁶. For the present study, scores were dichotomized by this clinically relevant split: (1) low depressive symptoms (scores ranging from 0 to 19) and (2) high depressive symptoms (scores 20 and above). According to this split, 42.7% of the sample reported high depressive symptoms.

Analytic Plan

Analyses were conducted in a series of steps. First, we determined the dimensionality of AUDs. In other words, through factor analysis, we explored whether the items loaded onto one single factor, indicating that AA and AD actually describe the same construct (presumably at different levels of severity), or loaded onto two separate factors (or categories) representing two distinct disorders. Since the primary statistical methodologies used to examine the dimensional versus categorical structure of AUDs are exploratory factor analysis (EFA) and confirmatory factor analysis (CFA)^{6,38}, an EFA was performed first on the 11 AUD criteria. Based on results from the EFA, a CFA was performed. The following guidelines proposed by Hu and Bentler³⁹ were used to assess for model fit in the CFA: Comparative Fit Index (CFI) > 0.95, Tucker-Lewis Index (TLI) > 0.95, and a root mean square error of approximation (RMSEA) < 0.06. We used MPlus software to conduct the CFA, which uses a robust unweighted least squares estimation to derive parameter estimates⁴⁰. For these analyses, tetrachoric correlations were specified because the AUD criteria were dichotomous.

Due to the fact that factor analysis does not provide information about how each criterion relate to one another in regards to item-level severity, IRT was used to explore item level characteristics. Specifically, IRT was used to determine (1) the rank ordering of each item's severity across the AUD continuum; (2) how well each item reliably classifies an individual with high or low severity of AUD; and (3) the functioning of each criterion to contribute additional information about the underlying severity of AUD (e.g., how important is one item compared to another in determining severity). Due to the fact that the CFA confirmed a two factor solution (described further below), thus violating the assumption of "unidimensionality" needed for IRT analysis on a single factor solution, IRT analysis was conducted independently on each set of the criteria (one analysis for AA items and one analysis for AD items). Two-parameter logistic models were applied separately to the four AA and seven AD SCID criteria, which provide estimates of both item *location* parameters (with values that typically range between -3 to 3) and the slope or item *discrimination* parameters (with values that typically range between 0 to 3). Larger location parameters indicate that higher values of the underlying latent-trait are necessary in order to endorse the item. In this case, a higher value for a location parameter would indicate greater severity of

AUD for that specific AUD criterion. In addition, larger discrimination parameters indicate that the probability of endorsing an item increases more rapidly as the underlying latent trait increases. In other words, a higher discrimination parameter for an AD criterion item would indicate that it functions well at classifying individuals as having AUD at varying levels of AUD severity.

Next, item characteristic curves (ICCs) and item information curves were generated for each of the AUD criteria. 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. These are used to determine which diagnostic criteria are more likely to be endorsed as the AUD severity increases. Item information curves indicate the point along the latent-trait continuum that an item is most reliable and are directly proportional to the magnitude of the discrimination parameter⁴¹. Total information curves were generated next—one for abuse and one for dependence. These curves were estimated by all values from the location and discrimination parameters for each item in a particular diagnosis, indicating the amount of information (i.e., reliability) the aggregate set of criteria provide across the underlying latent-trait AUD continuum.

Finally, exploratory analyses examined differential item functioning (DIF) across the location parameter in the following sub-groups: age (mean split into young vs. old), presence or absence of a SUD, and absence or presence of moderate to severe depression. These characteristics of participants were explored due the particular risk each of them poses for women⁴². For example, the high proportion of depression found among TANF recipients may influence the responses to AUD criteria. In addition, prior research has indicated significant age-related DIF across the AUD criteria in a general population of drinkers¹⁰. Due to the relatively low level of endorsement of AA criteria, only AD criteria were used for the DIF analyses.

All IRT models were analyzed using Multilog⁷⁴³, which estimates item parameters via a Bayesian expectation-maximization (EM) equation. The initial starting solution was set to assume a normal distribution for the latent-trait scores and, during each of the iterations; the posterior distribution from the prior EM was utilized in order to provide accurate estimation. In addition, the convergence criterion for the EM equation was set to .001.

RESULTS

Overall, the percent endorsement for each DSM-IV criteria ranged from 41.9% “drink more than intended (dependence)” to 1.1% “neglected role obligations (abuse)” (see Table 1). AD criteria were more highly endorsed than AA criteria. Approximately 41.2% of the sample met criteria for AD. About 4% met criteria for AA. EFA indicated that a two-factor model provided adequate fit to the AUD criteria (Table 1). The first factor accounted for approximately 56.1% of the common variance (eigenvalue = 6.18) with strong factor loadings for AD criteria (ranging from .952 to .881). The second factor accounted for approximately 29.5% of the common variance (eigenvalue = 3.24) with strong factor loadings for AA criteria (ranging from .985 to .788). The correlation between the factors was low ($r = .193$) suggesting a hierarchical order between factors. CFA confirmed this two-factor solution, which indicated good model fit (RMSEA = .035; TLI = .998; CFI = .999) over a one-factor solution.

IRT analyses further clarified the potential hierarchical relationship between the two factors via location and discrimination parameters (Table 1). The values of the location parameters for AD criteria were lower than the AA criteria and ranged from 0.26 to 0.75 with items “withdrawal”, “important activities given up”, and “recurrent physical/psychological

problems” providing the highest values. The location parameter values for AA criteria ranged from 2.25 to 2.77, with “neglected role obligations” providing the highest value. Overall, our findings imply that greater levels of the latent trait of AUD severity are necessary in order to respond to AA items than AD items. Additionally, values of the discrimination parameters ranged from 3.79 to 6.08 among AD items, with “drinking more than intended” and “activities to obtain alcohol” having the greatest overall discriminatory ability across the latent AUD severity continuum. Among AA items, discrimination parameters ranged from 2.78 to 6.06 with “drinking in hazardous situations” demonstrating the greatest discriminatory ability across the latent AUD severity continuum. The ICCs (Figure 1) illustrate this hierarchical relationship pictorially. Overall, each criterion indicates good discrimination across the underlying latent-trait AUD continuum, with AA criteria at the more severe end of the spectrum.

As noted by Neal and colleagues⁴⁴, when a total information curve is generally peaked it indicates the highest degree of reliability the scale has at that level of the underlying latent trait score. Figure 2 illustrates the item information curves. AD item information curves indicated high reliability at the lower end of the severity spectrum. In a separate analysis, AA items demonstrated lower reliability compared to AD items and fell at the higher end of the severity spectrum. Figure 3 illustrates the total information curves of AD and AA, respectively. The aggregate of the AD items demonstrated greater reliability than AA items at a lower end of severity. The aggregate of the AA items demonstrated a lower reliability than AD items at the higher end of severity.

For the purposes of this study, differential item functioning (DIF) analysis provided an examination of how the item content of the AD criteria only may be biased based on age (using a mean split), presence of an additional SUD, and presence of high depressive symptoms. Differences between groups with respect to DIF were tested using chi-square analyses. There were no significant DIF on the location parameters for age or presence of an additional SUD. With respect to depressive symptoms, significant DIF on the location parameters occurred for AD item “inability to cut down or stop” ($\chi^2(1) = 5.248, p < .05$). This finding indicates that the severity parameter was lower (46.6%; Location parameter = .07, SE = .12) for those with moderate to severe depression compared to those who did not have moderate to severe depression (36%; location parameter = -.227, SE = .101). This implies that at equivalent levels of AUD severity TANF recipients with high depressive symptoms were more likely to endorse this item than those with low or no depressive symptoms.

DISCUSSION

The present study extended existing research by examining the dimensionality and the internal structure of DSM-IV AUD criteria in a sample of TANF eligible women using techniques from factor analysis and IRT. A strong two factor solution for the AUD criteria was found, reflecting a separate factor for abuse and dependence, respectively. The AA criteria reflect an independent, separate factor that appears on the more severe end of the AUD severity continuum than AD criteria and show high discrimination. For this sample of TANF women, AA criteria indicate greater severity of AUD than AD criteria. Item information curves reveal that each of the items maintains its position and demonstrates high reliability in identifying individuals at a particular point of severity—the low end for AD criteria and the high end for AA criteria. The total information curves demonstrate that abuse criteria are more reliable than the dependence items at the more severe end of the spectrum. Based on these findings alone, maintaining two separate diagnoses of AA and AD are recommended.

These findings are inconsistent with existing literature on the dimensionality of DSM-IV AUD criteria which have found a strong, dominant one factor solution. Studies on the general population and other non-clinical populations demonstrate that AA and AD criteria are located on a continuum of AUD severity in an integrated or mixed fashion. This is contrary to our findings which found a strong two factor solution. Our findings may suggest that a two factor solution only emerges for individuals with the most severe clinical presentations of chemical dependency and high rates of co-occurring disorders.

The proposed modifications to AUD for DSM-5 include severity specifiers. Currently, those specifiers are based on symptom or criterion counts. Research suggests thus far that symptom counts, which are convenient and parsimonious, perform as well as potentially weighting criteria based IRT analyses to determine severity¹¹. Despite this empirical support, it may be that in the context of a complex population, such as TANF recipients, specific AUD criteria are better and stronger indicators of severity, and may provide additional information about an individual that could further elucidate the severity specifiers proposed for DSM-5. In this study, AA criteria reliably discriminated high severity of AUD. Even in studies where a one factor solution was strongly supported, some AA criteria (e.g., “role impairment”) are consistently located at the higher end of severity than the majority of the AD criteria^{8,10,19}. These findings point to the potential importance of attention to particular criteria as indicators of greater severity among specific populations.

This analysis underscores the importance of understanding measures in the context of multiple disorders. In this case, having moderate to severe depression increased the likelihood of endorsing the AD criteria “inability to stop or cut down.” This may represent a self-medication motive on the part the depressed. A clinician examining endorsement of this criterion among an individual with moderate to severe depression might interpret this as greater severity of AUD, as this criterion can indicate greater physical dependence; however, the depression itself may be distorting this severity. The importance of understanding how these criteria operate in the presence of co-morbid disorders (e.g., anxiety) is a crucial arena for future research.

This analysis provides important information that begins to fill current gaps in the literature on dimensionality of AUD criteria. First, this is the first analysis of its kind to specifically focus on a more clinically vulnerable population of women and ethnic minorities. Second, Saha and colleagues¹⁰ advocated for more research with different criterion measures for replication purposes, as different measures cull for item endorsement in different ways. The majority of the studies examining dimensionality of AUD have utilized the Alcohol Use Disorders and Associated Disabilities Interview Schedule^{10,11,13} or the Composite International Diagnostic Interview—Substance Abuse Module^{9,16,45}. Only one study has utilized the SCID and that focused only on adolescents⁸. This is the first study to use the SCID to measure AA and AD criteria on an adult population and therefore contributes important information to a collective of knowledge about dimensionality of AUD criteria. Third, Lagenbucher and colleagues¹⁹ called for the identification of criteria that indicate more severe cases. For women on TANF, such indicators appear to be AA criteria. Finally, to our knowledge, this is the first study to present any significant findings regarding the potential bias co-morbid mental health disorders may contribute in diagnosing AUD. The findings point to the importance of further exploration of the potential differences in criterion endorsement for those with co-morbid mental health disorders.

While findings from this study support a two factor solution to DSM-IV AA and AD criteria, it is important to consider these findings in the context of a preponderance of research that suggests the opposite. The current state of the research proposing a one-factor solution with symptom count severity specifiers does not take into account the unique

aspects of severity among special populations. Diagnoses are meant to identify and label the presence of illness, and in doing so, suggest a course of action for treatment. Among TANF women, a one-factor solution may in fact misrepresent the presentation of the illness and the specific aspects that identify a case as severe or not. This could potentially interfere with a referral to the appropriate level of care. Proceeding with modifications of DSM-IV AUD criteria should be done with caution.

Even if an integration of AA and AD proposed for DSM-5 is realized, it is important to note that the findings of this study are still pertinent because their clinical applications are important beyond specific diagnosis. Clinicians can gain a better understanding of each of their patients and their respective addictions by understanding how items are distributed along a latent trait severity continuum among different populations. In this case, clinicians working with TANF women need to be informed about which criteria may help to determine higher AUD severity (e.g., AA criteria) or potentially bias responses to certain items (e.g., AD criterion “cut down or quit”). This can provide nuanced information beyond a symptom count and diagnostic severity indicator that can impact treatment and subsequent prognosis. For example, it might help to determine that an individual needs a higher level of care of substance abuse treatment or indicate that an individual needs integrated mental health and AUD treatment. Clearly, a clinical evaluation of an individual differs from a research evaluation that asks about information in a rigid or inflexible manner; clinicians are best positioned to understand the nuances of an individual’s psychiatric state, as well as the unique aspects of his or her addiction and its severity beyond what is collected in a standardized measure. It is important to note, however, that social service systems are increasingly encouraged to adopt standardized measures to evaluate individuals, e.g.,^{46,47}, and reliance on such measures to dictate services and treatment will likely increase. Understanding how to interpret the findings of such measures allows for the determination of the best interventions for vulnerable populations.

Without more research on clinically complex populations, it is difficult to determine the exact clinical implications of the two factor model among TANF recipients. It is possible that the findings reflect the general struggle across the mental health field to differentiate primary and secondary SUD and mental health disorders. Those with high rates of multiple SUDs or co-occurring mental health disorders may endorse diagnostic items differently than other populations—in fact this is expected. The inability in this study to examine DIF on AA criteria prevents us from knowing more about how co-occurring disorders may be impacting responses to items on standardized measures such as the SCID. Still, the findings point to the importance paying attention to *all* the information provided by an instrument to inform a clinical evaluation, not just the total score or diagnosis. Though mental health and addiction treatment programs are becoming more adept at treating co-occurring disorders by treating them simultaneously, paying attention to individual responses to items on standardized measures may provide important information on for treatment plans of individual patients.

Limitations

There were several limitations associated with the current study. The most important limitation in this study is that almost all the women who endorsed AA criteria had a primary dependence diagnosis for another drug. As such, it could be that AA criteria in this case are simply a marker for other drug dependence—thus explaining the presence of these items at the extreme end of severity; however, given that so many of those with AD also had other drug dependence, it is difficult to say the exact nature of the influence of these multiple dependencies on these analyses.

Second, due to the low endorsement ratings for each of the AA criterion, the software program is not able to reliably (i.e., convergence criteria not be met) examine DIF across

each of these criterion. DIF could therefore only be explored in the context of AD. Third, results are generalizable to only the English speaking TANF recipients of Essex and Atlantic Counties, NJ or demographically similar welfare populations. Other state welfare populations are known to be quite demographically distinct. Fourth, the strength of these analyses is entirely dependent on the implementation of the SCID. While training and supervision of research staff was intensive, results are still reliant on the consistency across individual interviewers and truthful reporting by participants. Due to the fact that participants were mothers, their answers may have been socially biased out of fear of potential child welfare involvement or shame related to role obligation. Finally, the questions for AA criteria themselves may not operate well given the population in question. For example, individuals may inherently put themselves in dangerous situations by living in impoverished and high crime areas or role obligations for work may already be interrupted due to unemployment for a variety of reasons, of which substance use may be just one. Therefore these questions may not be a clear barometer of alcohol or other drug problems per se. As a result of all these limitations, results should be interpreted with caution.

Future Research

Findings highlight the need for further exploration of the dimensionality of AUD criteria for other vulnerable populations to move forward with the next edition of the DSM. This will not only continue to elucidate the specificity and reliability of the criteria across populations, but it will provide additional information and tools about which criteria may signal the most severe pathology. Additional study of the effects of co-morbid mental health disorders is also needed, as it may interfere with the diagnosis process and limit information and subsequent treatment for an individual's disorder.

Conclusion

For more than 20 years, experts in the substance abuse treatment field have advocated for a shift to a single construct of AUD via an integration of the two disorders into a single diagnosis e.g.,². While this study supports a continued two factor solution, based on the findings from previous studies described above, a one factor solution may still be the best, most parsimonious option for future AUD criteria and classification. These findings suggest, however, further thought must be given to the structure and utilization of the severity indicators. Finally, these findings provide a reminder of the importance of repeating analyses across populations to maximize our understanding and perspectives of AUDs.

Acknowledgments

This study was supported with funding from grant 5 R01 DA12256 from the National Institute on Drug Abuse, Bethesda, MD (Dr. Morgenstern); grant 90XP002 from the Administration for Children and Families, the Assistant Secretary for Planning and Evaluation; and the New Jersey Department of Human Services.

References

1. Edwards G, Gross M. Alcohol dependence: Provisional description of a clinical syndrome. *British Medical Journal*. 1976; 1:1058–1061. [PubMed: 773501]
2. Rounsaville BJ, Spitzer RL, Williams JB. Proposed changes in DSM-III substance use disorders: Description and rationale. *Am J Psychiat*. 1986; 143:463–468. [PubMed: 3953889]
3. Hasin D, Beseler CL. Dimensionality of lifetime alcohol abuse, dependence, and binge drinking. *Drug Alcohol Depen*. 2009; 101:53–61.
4. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4th ed.. Washington, DC: American Psychiatric Association; 2000. text revision ed.

5. 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]
6. Grant BF, Harford TC, Muthen BO, Hsiao-Ye YE, Hasin D, Stinson FS. DSM-IV alcohol dependence and abuse: Further evidence of validity in the general population. *Drug Alcohol Depen*. 2007; 86:154–166.
7. 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. *Journal of Studies on Alcohol*. 2001; 62:150–157. [PubMed: 11327181]
8. 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. *Journal of Abnormal Psychology*. 2006; 115(4):807–814. [PubMed: 17100538]
9. Krueger RF, Nichol PE, Hicks BM, et al. Using latent trait modeling to conceptualize and alcohol problems continuum. *Psychological Assessment*. 2004; 16(2):107–119. [PubMed: 15222807]
10. Saha TD, Chou SP, Grant BF. Toward an alcohol use disorder continuum using item response theory: Results from the National Epidemiological Survey on Alcohol and Related Conditions. *Psychological Medicine*. 2006; 36:931–941. [PubMed: 16563205]
11. 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 Depen*. 2010; 107:31–38.
12. Lynskey M, Nelson EC, Neuman RJ, et al. Limitations of DSM-IV operationalizations of alcohol abuse and dependence in a sample of Australian twins. *Twin Research and Human Genetics*. 2005; 8:574–584. [PubMed: 16354499]
13. 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. *Alcoholism: Clinical and Experimental Research*. 2006; 30(7):1165–1175.
14. American Psychiatric Association. DSM-5 development: Substance use and addictive disorders. 2010. <http://www.dsm5.org/ProposedRevision/Pages/SubstanceUseandAddictiveDisorders.aspx>. [Accessed August 8, 2011]
15. Kahler CW, Strong DR, Hayaki J, Ramsey SE, Brown RA. An item response analysis of the Alcohol Dependence Scale in treatment-seeking alcoholics. *J Stud Alcohol*. 2003; 64:127–136. [PubMed: 12608493]
16. Gelhorn H, Hartman C, Sakai J, et al. 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*. 2008; 47(11):1329–1339. [PubMed: 18827724]
17. Beseler CL, Taylor LA, Leeman RF. An Item Response Theory Analysis of DSM-IV alcohol use disorder criteria and "binge" drinking in undergraduates. *J Stud Alcohol Drugs*. 2010; 2010(71): 418–423. [PubMed: 20409436]
18. Hagman BT, Cohn AM. Toward DSM-V: Mapping the alcohol use disorder continuum in college students. *Drug Alcohol Depen*. in press.
19. Langenbucher JW, Labouvie E, Martin CS, Sanjuan PM, Bavly L, Kirisci L. An application of Item Response Theory analysis to alcohol, cannabis, and cocaine criteria in DSM-IV. *Journal of Abnormal Psychology*. 2004; 113(1):72–80. [PubMed: 14992659]
20. Kuerbis AN, Neighbors CJ, Morgenstern J. Depression's moderation of the effectiveness of intensive case management with substance dependent women on Temporary Assistance for Needy Families: Outpatient substance use disorder treatment utilization and outcomes. *J Stud Alcohol Drugs*. 2011; 72:297–307. [PubMed: 21388603]
21. Pollack HA, Reuter P. Welfare receipt and substance-abuse treatment among low-income mothers: The impact of welfare reform. *American Journal of Public Health*. 2006; 96(11):2024–2031. [PubMed: 17018836]
22. Dooley D, Prause J. Mental health and welfare transitions: Depression and alcohol abuse in AFDC women. *American Journal of Community Psychology*. 2002 Dec; 30(6):787–813. [PubMed: 12385483]

23. Rosen D, Tolman RM, Warner LA. Low-income women's use of substance abuse and mental health services. *Journal of Health Care for the Poor and Underserved*. 2004; 15:206–219. [PubMed: 15253374]
24. Brems C, Johnson ME. Clinical implications of the co-occurrence of substance use and other psychiatric disorders. *Professional Psychology: Research and Practice*. 1997; 28(5):437–447.
25. Watkins KE, Hunter SB, Wenzel SL, et al. Prevalence and characteristics of clients with co-occurring disorders in outpatient substance abuse treatment. *Am J Drug Alcohol Ab*. 2004; 30(4): 749–765.
26. Besharov, D.; Germanis, P. Full-engagement welfare in New York City: Lessons for TANF's participation requirements. Washington, DC: American Enterprise Institute, Welfare Reform Academy; 2004 Aug.
27. Gallagher, L.J.; Uccello, CE.; Pierce, AB.; Reidy, EB. State general assistance programs 1998. Washington, DC: The Urban Institute; 1999 Apr.
28. Morgenstern J, Hogue AT, Dasaro C, Kuerbis A, Dauber S. Characteristics of individuals screening positive for substance abuse in a welfare setting: Implications for welfare and SUD treatment systems. *J Stud Alcohol Drugs*. 2008; 69:561–570. [PubMed: 18612572]
29. First, MB.; Spitzer, RL.; Gibbon, M. Structured Clinical Interview for DSM-IV. New York: Biometric Department, New York State Psychiatric Institute; 1996.
30. Morgenstern J, Blanchard K, McCrady BS, McVeigh KH, Morgan TJ, Pandina RJ. Effectiveness of intensive case management for substance dependent women receiving Temporary Aid to Needy Families (TANF). *American Journal of Public Health*. 2006; 96(11):2016–2023. [PubMed: 17018819]
31. Morgenstern J, McCrady BS, Blanchard KA, McVeigh KH, Riordan A, Irwin TW. Barriers to employability among substance dependent and non-substance-affected women on federal welfare: Implications for program design. *J Stud Alcohol*. 2003; 64(2):239–246. [PubMed: 12713198]
32. Morgenstern, J.; Riordan, A.; Dephillippis, D., et al. Specialized screening approaches can substantially increase the identification of substance abuse problems among welfare recipients. Washington, DC: Administration for Children and Families; 2001.
33. Center for Substance Abuse Treatment. Supplementary administration manual for the expanded female version of the Addiction Severity Index (ASI) instrument: The ASI-F. Washington, DC: Government Printing Office; 1997.
34. Makela K. Studies of reliability and validity of the Addiction Severity Index. *Addiction*. 2004; 99:398–410. [PubMed: 15049734]
35. Beck, AT.; Steer, RA.; Brown, GK. Beck Depression Inventory, Second Edition Manual. San Diego, CA: Harcourt Brace; 1996.
36. Greenfield SF, Weiss RD, Muenz LR, et al. The effect of depression on return to drinking: A prospective study. *Archives of General Psychiatry*. 1998; 55:259–265. [PubMed: 9510220]
37. Dum M, Pickren J, Sobell LC, Sobell MB. Comparing the BDI-II and the PHQ-9 with outpatient substance abusers. *Addict Behav*. 2008; 33:381–387. [PubMed: 17964079]
38. 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*. 2001; 62(2):150–157. [PubMed: 11327181]
39. Hu, LT.; Bentler, P. Evaluating model fit. In: Hoyle, RH., editor. *Structural equation modeling: Concepts, issues, and applications*. London: Sage; 1995. p. 76-99.
40. Muthen, L.; Muthen, BO. *Mplus user's guide*. Los Angeles, CA: Muthen and Muthen; 1998–2001.
41. Embretson, SE.; Reise, SP. *Item Response Theory for Psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.; 2000.
42. Green CA, Polen MR, Dickinson DM, Lynch FL, Bennett MD. Gender differences in predictors of initiation, retention, and completion in an HMO-based substance abuse treatment program. *Journal of Substance Abuse Treatment*. 2002; 23:285–295. [PubMed: 12495790]
43. Thissen, D.; Chen, WH.; Bock, D. *Multilog software*. Lincolnwood, IL: Scientific Software International; 2010.

44. Neal DJ, Corbin WR, Fromme K. Measurement of alcohol-related consequences among high school and college students: Application of item response models to the rutgers alcohol problem index. *Psychological Assessment*. 2006; 18(4):402–414.
45. World Health Organization. *Composite International Diagnostic Interview (CIDI)*. Geneva, Switzerland: World Health Organization; 1990. Version 1.0 ed.
46. Berkman B, Maramaldi P. Use of standardized measures in agency based research and practice. *Social Work in Health Care*. 2001; 34(1–2):115–129. [PubMed: 12219762]
47. Buchanan A, Ritchie C. Using standardized measures to prioritize services for children and families in need. *Journal of Social Work*. 2004; 4(2):167–178.

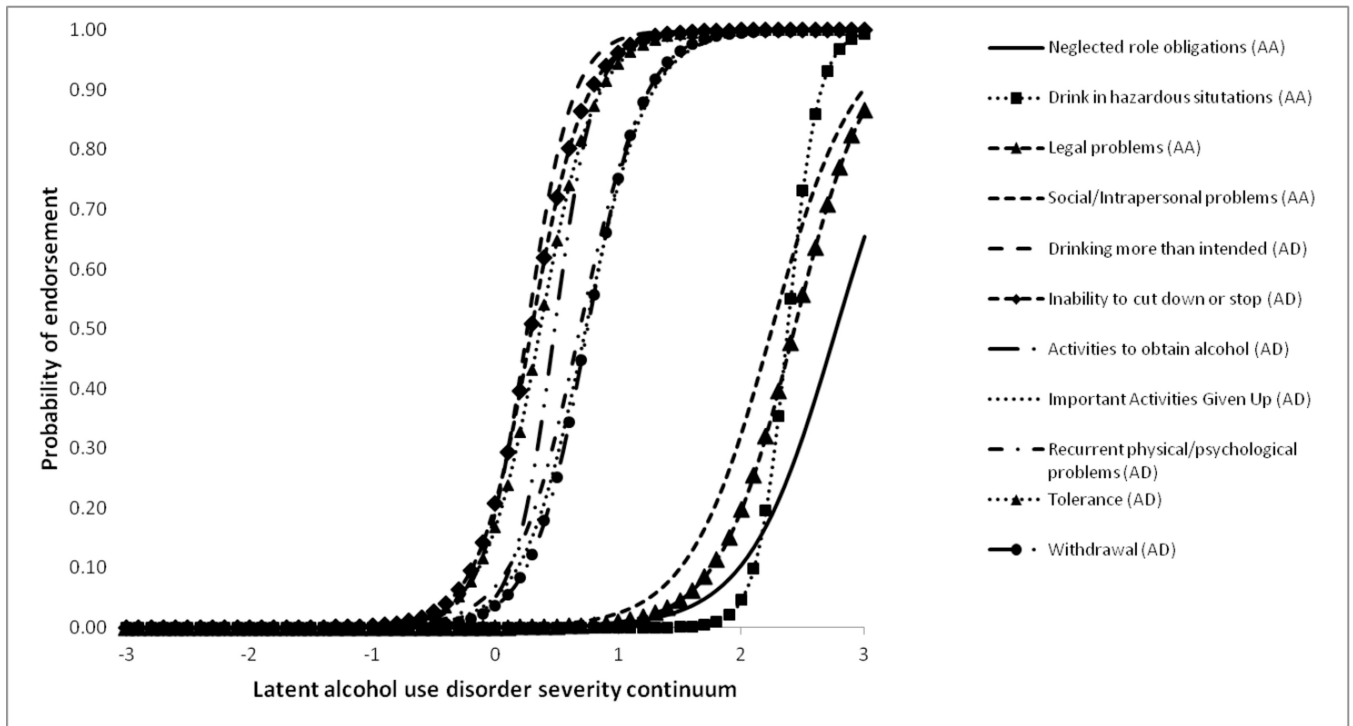


Figure 1.
Item characteristic curves by criteria abuse and dependence.

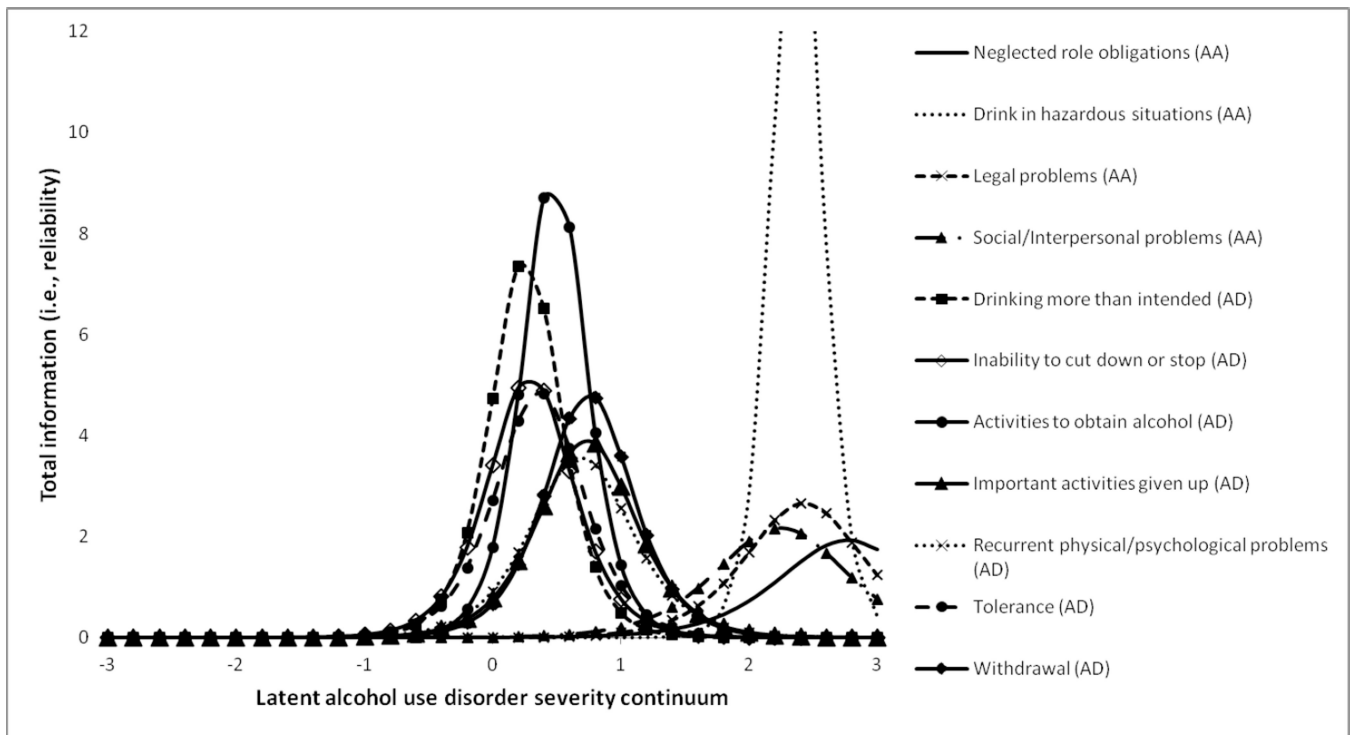


Figure 2.
Item information curves for alcohol abuse (AA) and alcohol dependence (AD) criteria.

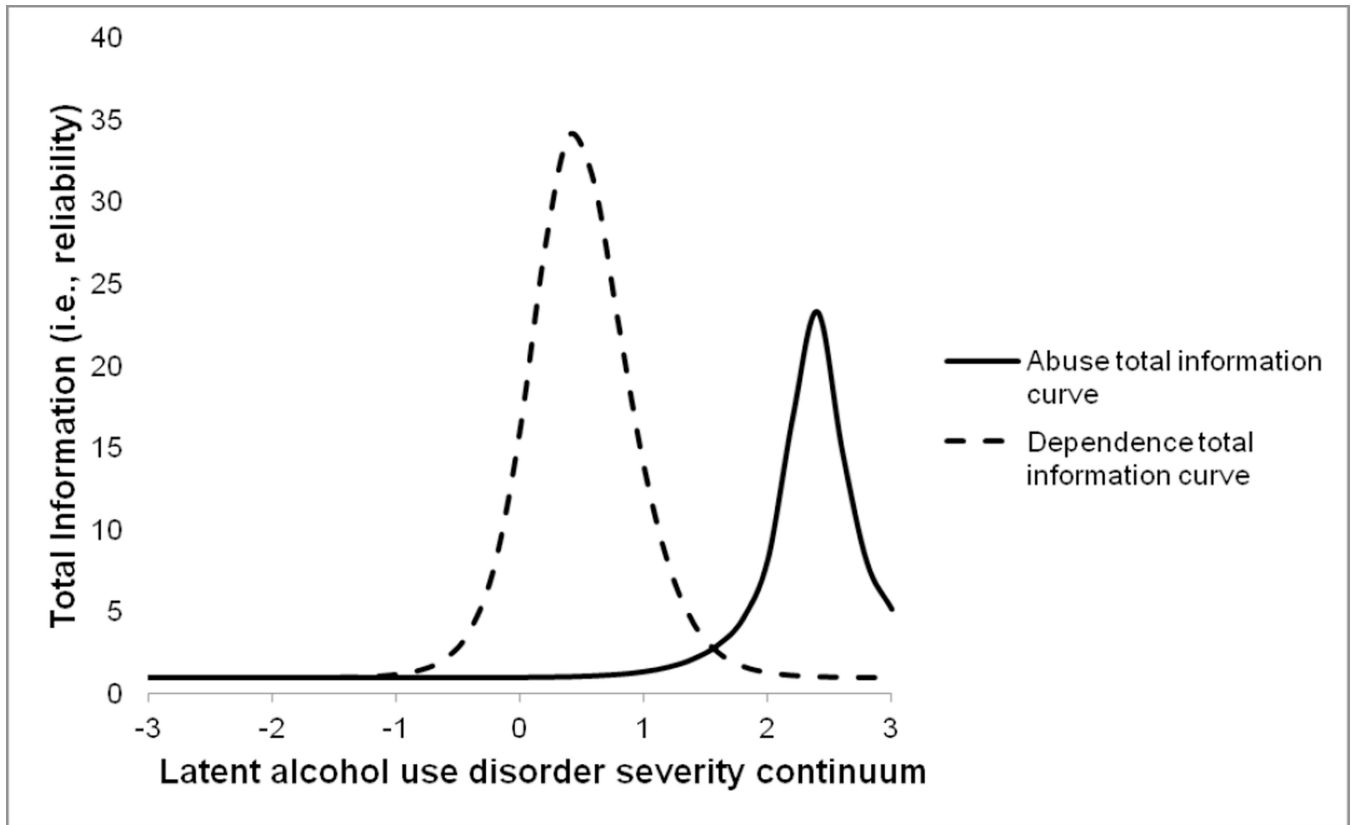


Figure 3.
Total information curve for alcohol abuse and dependence

Exploratory and confirmatory factor analysis loadings and location and discrimination parameters of the criteria for alcohol abuse and dependence

Table 1

	% Endorsed	EFA 1	EFA	Location	SE	Discrimination	SE
Drinking more than intended (AD)	41.9	0.952	0.212	0.26	0.05	5.49	0.74
Inability to cut down or stop (AD)	40.3	0.925	-0.062	0.29	0.05	4.55	0.59
Activities to obtain alcohol (AD)	33.2	0.941	-0.191	0.48	0.04	6.08	1.15
Important activities given up (AD)	24.5	0.885	-0.197	0.73	0.06	3.96	0.71
Recurrent physical/psychological problems (AD)	26.2	0.881	-0.269	0.69	0.06	3.79	0.65
Tolerance (AD)	37.7	0.916	-0.072	0.36	0.05	4.41	0.63
Withdrawal (AD)	23.6	0.897	-0.229	0.75	0.05	4.39	0.75
Neglected role obligations (AA)	1.1	-0.183	0.788	2.77	0.41	2.78	1.02
Drink in hazardous situations (AA)	1.1	-0.191	0.985	2.38	0.14	6.06	2.83
Legal problems (AA)	1.7	-0.291	0.843	2.43	0.24	3.27	0.94
Social/Interpersonal problems (AA)	2.8	-0.142	0.841	2.25	0.23	2.95	1.21
Eigenvalue		6.176	3.24				
% variance accounted for		56.14	29.54				
		%	%				
RMSEA (CFA)		0.129	0.035				
CFI (CFA)		0.979	0.999				
TLI (CFA)		0.973	0.998				