Published in final edited form as:

Alcohol Clin Exp Res. 2016 August ; 40(8): 1691–1699. doi:10.1111/acer.13121.

Assessment of Withdrawal and Hangover are Confounded in the Alcohol Use Disorder and Associated Disabilities Interview Schedule: Withdrawal Prevalence is Likely Inflated

Cassandra L. Boness, Sean P. Lane, and Kenneth J. Sher

University of Missouri

Abstract

Background—The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV) and -5 are diagnostic interviews used in major epidemiological and other studies of alcohol use disorder (AUD). Much of what we know regarding the prevalence of AUD in the United States is based upon this interview. However, past research and meta-analytic evidence suggest that differential operationalization of the AUD criteria across instruments can lead to differential endorsement of symptoms and resulting AUD diagnosis rates. In particular, studies employing the AUDADIS are observed to have markedly higher endorsement rates of withdrawal than other large epidemiological studies. One explanation for this is that when assessing withdrawal, the AUDADIS combines effects from the morning after drinking with those from the days following, thereby conflating hangover and withdrawal.

Methods—The current study addresses whether this operationalization confounds rates of endorsement when compared to simpler, less ambiguous hangover or withdrawal stems. To this aim, 497 college student drinkers were randomized into one of three stem conditions: (1) hangover ($n = 164$), (2) withdrawal ($n = 167$), or (3) combined AUDADIS-IV ($n = 166$).

Results—Across conditions, participants were more likely to report the occurrence of each withdrawal symptom in the combined stem condition than in the explicit withdrawal stem condition, but not in the explicit hangover stem condition. Within the combined stem condition, probed symptoms were more likely to be reported as a result of a hangover.

Conclusion—The AUDADIS potentially results in false positives for withdrawal, arguably a pathognomic symptom of alcoholism and, in turn, likely affects rates of the diagnosis of AUD.

Keywords

Alcohol Use Disorder and Associated Disabilities Interview Schedule; alcohol use disorder; assessment; diagnosis; alcohol withdrawal

Correspondence concerning this article should be addressed to Cassandra Boness, Department of Psychological Science, Psychology Building, 200 South Seventh Street, Columbia, MO 65211. *Phone*: +1 573 884 1485. *Fax*: +1 573 884 5588. clmkdb@mail.missouri.edu.

Author Note: Cassandra L. Boness, Sean P. Lane, and Kenneth J. Sher, Department of Psychological Science, University of Missouri.

Introduction

The Alcohol Use Disorder and Associated Disabilities Interview Schedule – Fourth Edition (AUDADIS-IV; Grant Dawson & Hasin 2001) and Fifth Edition (AUDADIS-5; Grant et al. 2011) are fully structured diagnostic assessments designed to evaluate alcohol, drug, and mental disorders using DSM-IV and DSM-5 diagnostic criteria, respectively (American Psychiatric Association [APA] 2000; 2013). These assessments are intended for administration in clinical and general populations (Grant Dawson & Hasin 2001; Grant et al. 2011). The AUDADIS-IV and -5 have been used in the National Institute on Alcohol Abuse and Alcoholism's (NIAAA) National Longitudinal Alcohol Epidemiologic Survey (NLAES; Grant Peterson Dawson & Chou 1994), in Waves I and II of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC; Grant & Kaplan 2005; Grant Moore & Kaplan 2003), and in NESARC-III (Grant et al. 2015a). Consequently, much of what we know regarding the prevalence of alcohol use disorders (AUD) in the United States is based upon this instrument.

The reliability and validity of the AUDADIS-IV has been evaluated in clinical (Canino et al. 1999; Hasin et al. 1997; Ruan et al. 2008) and general populations (Grant et al. 1995; 2003), both in the United States and in other countries (Chatterji et al. 1997; Compton Thomas Stinson & Grant 2007; Vrasti et al. 1997). Similarly, the AUDADIS-5 has demonstrated good to excellent test-retest reliability for substance use disorder diagnoses in the general population (Grant et al. 2015b). The AUDADIS-IV and 5 also demonstrate procedural “validity,” based on its concordance with the Psychiatric Research Interview for Substance and Mental Disorders, DSM-5 (PRISM-5; Hasin et al. 2015). Overall, the AUDADIS-IV and -5 appear to have adequate reliability and validity across a variety of samples when assessed in these traditional ways.

However, in a recent meta-analysis, Lane et al. (in press) demonstrated, from an item-response theory (IRT; Embretson & Reise 2000) perspective, that a number of AUD criteria are systematically differentially severe in the AUDADIS compared to other assessments. One criterion that has previously received considerable attention, but for more substantive reasons, is withdrawal (Caetano & Babor 2006; Martin et al. 2008; 2011; 2014). Specifically, Lane and colleagues (in press) found that, independent of previously studied factors such as age, gender, and clinical/general samples, the relative severity of withdrawal in NESARC is systematically lower than the relative severity of withdrawal reported in other epidemiological studies (e.g., the National Survey on Drug Use and Health [NSDUH; Substance Abuse and Mental Health Services 2006] and large scale studies such as the Collaborative Studies on Genetics of Alcoholism [COGA; Bierut et al. 1998]). In fact, NESARC withdrawal severity was amongst the lowest of the 30 unique samples examined in the meta-analysis. For example, withdrawal was considered “highly” severe in the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) used in the COGA and other studies, and the Substance Abuse and Mental Health Services Administration (SAMHSA) survey used in NSDUH, but only “middling” in the AUDADIS-IV.

More concretely, past 12 month prevalence rates of withdrawal ranged from 7.8-8.2% in NESARC (Casey et al. 2012; Saha et al. 2006) and were 12.6% in NLAES (Keyes et al.

2011), both of which employed the AUDADIS, while analogous prevalence rates were considerably lower in population studies that used other assessments – 1.1-3.9% for the NSDUH interview (e.g. Harford et al. 2009) and 3.3% for the National Survey of Mental Health and Well-Being, which employed the Composite International Diagnostic Interview (e.g. Proudfoot et al. 2006). We cannot definitively say that these differences in withdrawal prevalence directly lead to parallel differences in overall AUD prevalence rates measured using the AUDADIS versus other assessments because of other factors such as current drinker status, differences in thresholds for symptom occurrence, past year versus lifetime diagnosis timeframes, and DSM-IV and DSM-5 diagnostic algorithms. However, epidemiological estimates of AUD indicate prevalences of approximately 7.4-12.7% for studies using the AUDADIS (Grant et al. 2004; Grant et al. 2015c) but 6.0-7.2% for studies using different assessments (Harford et al. 2009; Proudfoot et al. 2006), which is consistent with this withdrawal difference at least being a factor. Thus, examining the differences in how withdrawal is assessed across diagnostic instruments is of great importance.

One potential explanation for the discrepancy between instruments with regard to withdrawal severities is that in assessing symptoms of withdrawal, the AUDADIS begins the section on withdrawal with the preamble: “The next few questions are about the bad aftereffects of drinking that people may have when the effects of alcohol are wearing off. This includes the morning after drinking or in the first few days after stopping or cutting down. Did you ever...” (Grant et al. 2001). Consequently, this stem potentially conflates two phenomena: (a) hangover (i.e., “the morning after drinking”) and (b) withdrawal (i.e., “in the first few days after stopping or cutting down”). As a result, the AUDADIS withdrawal items likely measure hangover-related symptoms as well as withdrawal and, thus, 1) fail to distinguish the two and 2) potentially incorrectly designate hangover symptoms as withdrawal symptoms.¹

Although hangover and withdrawal share overlapping symptomatology (e.g., nausea, restlessness, depression, and anxiety) and hangover is sometimes conceptualized as a type of acute withdrawal (Newlin & Pretorius 1990; Prat et al. 2009), existing research indicates that hangover is physiologically and clinically distinct from alcohol withdrawal, the latter of which is a result of neuroadaptation after a pattern of long-term consumption (see Prat et al. 2009 for a more in-depth discussion). Arguably, an individual may have reported experiencing many of the symptoms associated with withdrawal, but in the context of a hangover experience (i.e., following acute intoxication), even if genuine withdrawal was never experienced. Subsequently, false positive assessments of withdrawal may be generated when the AUDADIS is used because it conflates withdrawal with hangover. The conflation of withdrawal with hangover might contribute to relatively low relative severity of withdrawal in the NESARC sample (Saha et al. 2006), in which withdrawal tends to be a moderate severity criterion, compared to samples like COGA (Bierut et al. 1998), in which withdrawal tends to be a high severity criterion (Lane et al. in press). Notably, in the COGA

¹From a measurement perspective, assessment instruments should include items that cover a range of severities, or difficulties, in order to adequately assess all levels of a given construct. However, including hangover items in an assessment intended to measure withdrawal will lower the severity estimate of the withdrawal criterion and decrease coverage over the focal part of the distribution (since hangover is a more common somatic experience), which can then unduly influence IRT and factor analyses.

study, the assessment instrument used, the SSAGA, takes special caution to exclude hangover experiences from withdrawal symptomatology. Specifically, SSAGA uses the stem:

“People who cut down, stop, or go without drinking after drinking steadily for some time may not feel well. These feelings are more intense and can last longer than the usual hangover. When you stopped, cut down, or went without drinking, did you ever experience any of the following problems for most of the day for 2 days or longer...” (Bucholz et al., 1994)

The SSAGA clearly makes an attempt to disambiguate hangover from withdrawal and avoid any potential confusion of the two phenomena. In contrast, the AUDADIS fails to draw a clear distinction between the two.

Although the AUDADIS-IV was developed with DSM-IV in mind, the DSM-IV does not use any verbiage that could be mistaken for hangover. The DSM-IV requires that alcohol withdrawal symptoms occur after, “Cessation of (or reduction in) alcohol use that has been heavy and prolonged” (APA 2000, p. 198). The DSM-5 (APA 2013) uses the same wording. Unlike the DSM-IV and DSM-5, the AUDADIS mentions, “the morning after drinking” and, as noted above, steps were not taken to ensure withdrawal and hangover are not conflated.

This is not, however, the first paper to suggest misidentification of AUD due to measurement error in general population surveys (Caetano & Babor, 2006; Midanik, Greenfield, & Bond, 2007). For example, Caetano and Babor demonstrated that in the 2001 NSDUH general population survey of those 12 years of age and older, symptom level prevalence estimates showed that younger age groups, when compared to older age groups, tend to report higher rates of tolerance and withdrawal symptoms. The authors hypothesize this may be related to the wording structure of the interview schedules which may conflate acute intoxication (e.g., binge or heavy drinking episodes) and the physical symptoms of alcohol withdrawal. Similarly, Harford and colleagues (2005), who found similar patterns in the 2001 National Household Survey on Drug Abuse, suggest the age differences in withdrawal prevalence may be related to a relative lack of experience with the effects of alcohol.

Whereas in the DSM-5 withdrawal is only one of the eleven criteria in the AUD criterion set, its importance is considerable for a number of reasons. First, some researchers argue that it is pathognomonic of alcohol dependence (i.e., it should be considered both necessary and sufficient for the diagnosis of alcoholism; Langenbucher et al. 2000). Second, withdrawal is associated with a kindling-like phenomenon that becomes progressively more severe with repeated withdrawals (Becker 1998; Breese, 2005). Finally, the presence of withdrawal has specific treatment indications (Bayard et al. 2004).

Here, we address whether the use of the “combined stem” of the AUDADIS yields different rates of endorsement as compared to a simple hangover or withdrawal stem. We hypothesized that, in a young adult sample, the combined stem would generate higher reported experience of withdrawal symptoms than would an explicit withdrawal stem. We expected hangover to be more common in a sample of young adult heavy drinkers, whose relative lack of drinking experience and binge-like drinking patterns would tend to yield

high rates of hangover and relatively lower rates of withdrawal, than in a sample of older adults. This is especially so if withdrawal is conceptualized as the neuroadaptation to a chronic drinking pattern. Because we expected the rates of hangover to be high in a collegiate sample (relative to withdrawal) due to their shorter drinking histories (relative to adults), we anticipated that the endorsement rates of symptoms preceded by the “pure” hangover stem would be relatively similar to those endorsed by the combined stem. We also included a “headache” item (which is included in the AUDADIS, but not scored as a symptom of withdrawal) for comparative purposes because “headache” is a classic hangover symptom, and not a part of the withdrawal item set.

Materials and Methods

Sample

Participants in the study were 526 college undergraduates enrolled in Introduction to Psychology courses at the University of Missouri (a large state university with approximately 20,000 undergraduates in attendance) during the Fall 2014 semester. Participants were recruited via an online research participation system. Participants self-selected to participate in the study. To be eligible for participation, students had to be over the age of 18 and have consumed alcohol on at least 12 separate occasions in the past 12 months. Twenty (3.8%) participants were excluded from analyses because they did not meet eligibility criteria (16 did not reach the alcohol consumption criterion, 3 did not consent to participate, 1 was under 18 years old). Nine additional participants (1.7%) started but did not complete the entire assessment, reflecting a response rate of 98.3%. These individuals were excluded from the analyses given the significant amount of missing data on the variables of interest. The total sample size comprised 497 students (mean age = 18.68; 41.85% male), which is an overall inclusion rate of 94.5%. Compensation for participation was awarded in the form of course credit. The 497 participants were randomly assigned to one of three conditions: (1) hangover stem ($n = 164$), (2) withdrawal stem ($n = 167$), or (3) combined AUDADIS-IV stem ($n = 166$). Participant demographics were similar across groups (see Table 1).

Measures

Items from the AUDADIS-IV (Grant Dawson & Hasin 2001) were used to assess DSM-IV withdrawal symptoms in the past 12 months via self-report using an online questionnaire. The AUDADIS-IV assesses 8 withdrawal symptoms (see Table 2 for a list of the withdrawal symptoms), and two relief drinking symptoms, consistent with the DSM-IV and DSM-5, which require two or more withdrawal symptoms to be endorsed, or one or more of the relief/avoidance drinking items to be endorsed. We operationalized endorsement as participants' indication that they had experienced the individual symptoms at least once in the past 12 months. For the purposes of this study, only the withdrawal symptoms were assessed because the relief drinking items do not employ the potentially problematic combined stem.

Additionally, all participants were assessed for the other 10 AUD symptoms employing a range of items. These items were derived and/or adapted from pre-existing diagnostic

interviews, self-report scales, and the DSM manuals. Participants also reported their alcohol use in the past year (frequency of drinking, frequency of getting drunk, and quantity consumed when drinking), and possible alcohol-related consequences. These additional items were included to address a different set of research goals related to criterion coverage and are not discussed further.

Procedure

The following procedures were approved by The University of Missouri Human Subjects Institutional Review Board. Participants who met the requirements for participation in the study were assigned an online time-slot during which they could complete the study. They were required to complete an approved informed consent form before beginning the questionnaire.

After completing the information on AUD symptoms, alcohol use, and alcohol-related consequences, participants were randomly assigned to one of three conditions: (1) the hangover stem (“In the past 12 months, after drinking alcohol, did you EVER find yourself...”), (2) the withdrawal stem (“In the past 12 months, when you stopped, quit, or cut down, did you EVER...”), or (3) the combined AUDADIS-IV stem (“The next few questions are about the bad aftereffects of drinking that people may have when the effects of alcohol are wearing off. This includes the morning after drinking or in the first few days after stopping or cutting down in the past 12 months. In the past 12 months, when the effects of alcohol were wearing off did you EVER...”). Consistent with standardized assessments, each condition then presented and assessed the same 9 withdrawal symptoms in a fixed order.

In the hangover and withdrawal stem conditions, participants reported the number of times in the past year they had experienced each symptom on a 5-point scale ranging from never to 4+ times in the past year. Symptoms must have occurred at least one time in the past year to qualify as “present” in the current analyses, which is consistent with the AUDADIS. In the combined stem, participants were first asked whether they had experienced any of the 9 symptoms in the past year. If they indicated they had experienced a specific symptom, participants answered probes regarding whether this symptom was experienced “the morning after drinking” (i.e., hangover), “after stopping, quitting, or cutting down” (i.e., withdrawal), or “both” on a yes/no scale. This design feature allowed for comparisons of withdrawal endorsement between the different conditions, as well as an internal replication within the combined stem condition itself.

Results

Differences across Conditions

Table 2 shows the reported endorsement rates associated with each of the stems and the associated chi-square tests of association for each symptom.² In order to set the experiment-wise error for detecting a stem effect on item-specific endorsement rates, a nominal p level

²Given that there were some participants that endorsed delirium tremens and/or seizures, which is highly unlikely in this sample, all analyses presented below were redone after those endorsing DTs and/or seizures were removed ($N=39$). The pattern of results

of 0.0056 (0.05/9) was set. This includes the headache item. Each of these nine tests were significant at this level. *Post hoc* pairwise differences across the three stems were assessed using logistic regression analyses. In the majority of analyses, there were significant differences in endorsement rates between the combined stem and the withdrawal stem, but not the combined stem and the hangover stem (see Table 2). Further, on average, participants in the hangover and combined stem conditions reported more symptoms on average (including headache). Thus, the rate of reporting of symptoms given the hangover stem appears to be more similar to the rate of reporting of symptoms in the the combined stem, both of which are much higher than in the withdrawal stem for the majority of symptoms assessed. This pattern of findings strongly suggests the combined stem may be producing false positives in assessment of withdrawal.

Within-Subject Comparisons: Probing the Reasons for Endorsing the Combined Stem

Recall that individuals that report experiencing symptoms (i.e., positively endorse the symptom) in the combined stem condition were further queried as to whether they had experienced the symptom “the morning after drinking” (i.e., hangover) and/or “after stopping, quitting, or cutting down” (i.e., withdrawal). Table 3 shows the base rate of combined stem endorsement and the overall implied endorsement rate of withdrawal and hangover (assuming that not endorsing the combined stem indicates neither symptoms of hangover nor withdrawal). McNemar tests (in which non-endorsement of the symptom implied negative endorsement of both hangover and withdrawal and was coded as such for analytic purposes) demonstrated that endorsement of the combined stem items was more strongly associated with reported hangover than with reported withdrawal, with the exception of the delirium tremens (DTs) and seizures symptoms, which were non-significant (see Table 3).

Discussion

Our findings are consistent with our hypothesis: The AUDADIS combined stem over-diagnoses withdrawal symptoms due to its conflation of hangover and withdrawal. This was evident in both between-subjects analyses yielding differential endorsements across stem conditions, and within-subjects analyses of the combined stem condition, indicating endorsement of the combined stem is more strongly related to hangover than withdrawal symptoms. The consistency of results between- and within-conditions provide a compelling internal replication of findings.

Further, these findings provide a likely explanation for why withdrawal appears to be only a moderate severity criterion in NESARC (Saha et al. 2006) as opposed to a high severity criterion in the SSAGA, in which the wording for withdrawal symptomatology is less open to interpretation (Bucholz et al. 1994). These findings are not meant to call into question other findings in NESARC, NESARC-III, or NLAES based on the AUDADIS. However, they do call into question the accuracy of prevalence estimates of DSM-IV dependence (APA 2000) and DSM-5 AUD (APA 2013) which are based, in part, on withdrawal

remained virtually identical. However, we believe it is important that these individuals are included in the primary analyses as their somewhat unlikely response patterns may be present in general population samples.

symptomatology. Moreover, some researchers view withdrawal as a cardinal symptom of dependence/addiction (Schuckit et al. 1998; Langenbucher et al. 2000) and Langenbucher et al.'s Withdrawal-Gate Model considers withdrawal pathognomonic (i.e., necessary and sufficient) of dependence. Researchers and clinicians planning to use the AUDADIS in their studies or practice would do well to consider supplementation of the AUDADIS' standard assessment of withdrawal or, alternatively, use an instrument with greater conceptual clarity for assessing withdrawal. With respect to taking likely over-diagnosis in published studies using the AUDADIS into account, estimates of withdrawal, in specific, and AUD, in general, are likely inflated and thus estimates from studies such as NESARC and NESARC-III should be viewed in that context.

To that end we conducted follow-up analyses using the NESARC-III dataset and the results of the current study to estimate how observed epidemiological prevalence estimates of AUD would be affected assuming only a portion of the NESARC-reported withdrawal endorsements were true positives. To do so we estimated past year AUD prevalence a) using the DSM-5 specifications, b) using the DSM-5 specifications but excluding withdrawal symptoms (and including relief/avoidance as true positives for withdrawal) to establish a lower bound, and c) using the odds ratios from Table 3 (i.e. within-person) to weight the probability of true endorsement given actual observed endorsement in NESARC-III for each of the 8 AUDADIS withdrawal symptoms within-person (including relief/avoidance endorsement as true positives).

We replicated the DSM-5 AUD prevalence reported by Grant and colleagues (Grant et al. 2015c) of 13.87% ($n = 5,133$) for NESARC-III. Excluding withdrawal symptoms from diagnosis (and keeping drinking for relief or avoidance of withdrawal) this drops to 13.12% ($n = 4,877$), a relative decrease of 5.0%. Weighting the probability of true endorsement for each withdrawal symptom given the odds ratios obtained from the current study, the prevalence is 13.14% ($n = 4,882$), a 4.9% decrease in prevalence relative to that reported by Grant et al. (2015c). If we restrict the sample to young adults (18-24 years; $n = 4,496$) DSM-5 AUD prevalence is 27.4% ($n = 1,159$), which drops to 25.6% ($n = 1,089$), a relative decrease of 6.0%, when the withdrawal criterion is either dropped or weighted. Furthermore, we note that the endorsement rate of the combined withdrawal criterion in NESARC-III is 7.7%, which is very similar to past studies using the AUDADIS (Casey et al. 2012; Saha et al. 2006), while our adjusted withdrawal criterion endorsement rate is 3.3%, which is very similar to that of assessments that take care to exclude hangover (e.g. Duncan et al. 2011; Harford et al. 2009; Proudfoot et al. 2006).

The results of the NESARC analyses and those of our primary study are consistent with a study by Karriker-Jaffe, Witbrodt, and Greenfield (2015), who found that of the 225 current drinkers in a sample of people under age 46 with at least 1 AUD symptom, 11% reported past-year withdrawal symptoms. Follow-up items suggested that 28% of those reporting past-year withdrawal symptoms indicated that the symptoms had occurred within 8 hours of stopping drinking. This suggests the participants may have been reporting the effects of acute intoxication instead of physical withdrawal symptoms, resulting in inflated estimates of withdrawal. The authors adjusted for this and estimated that it resulted in approximately a

6% inflation in AUD prevalence, very similar to the estimates of our follow-up analyses of the NESARC-III data.

Although the issue of withdrawal symptoms being confounded with hangover symptoms stands out, there are several additional potential issues with AUD diagnostic criteria, as well as with the instruments used to assess symptoms other than withdrawal that should be considered. For example, there are multiple operationalizations used to assess the same AUD criterion both within and across instruments (Martin & Chung 2008). These operationalizations may result in different rates of endorsement based on the severity of the item, which potentially impacts the relative severity of that criterion relative to the other 10 criteria. Further, the threshold, or number of times a given symptom must occur to be considered “present” often differs between operationalizations, criterion, and instruments (e.g. Grant et al. 1995; 2001; 2011; Bierut et al. 1998). Given the results of the present study and these additional issues, there is a need for a more systematic exploration of the potential limitations of AUD diagnostic assessments or, at least, greater appreciation of the different biases that different assessments possess.

These conclusions should be considered within the context of the study's limitations. Perhaps the major limitation of this study is the nature of sample (collegiate), limiting generalization to other types of populations. We note, however, that the college years represent the highest period of prevalence of AUD across the lifespan and, thus, this potential bias is likely to have a significant influence in overall prevalence estimates spanning adulthood (Grant et al. 2015c). The current study may also be limited by the fact that the AUDADIS is a structured interview typically administered in a face-to-face format whereas here, withdrawal was assessed using an online survey. Participants potentially responded differently in this study than they might in a face-to-face interview, although existing research has suggested diagnostic interviews administered by computer yield acceptable diagnostic validity and reliability (Blouin Perez & Blouin 1988; Miller et al. 2004; Skinner & Allen 1983) and in other assessment spheres seem to show equal validity compared to those administered under more controlled, face-to-face contexts (Gosling et al., 2004). There may be additional concern that participants responded differently to the items due to the items being presented with different stems, however, we consider this unlikely since the within-subject comparisons strongly replicate the between subject comparisons and these alternate approaches do not share similar experimental biases.

Further limitations include withdrawal symptoms within each condition not being presented in a random order between participants. Although this may have produced an order effect, we are confident in the robustness of our findings given these patterns occur both between and within subjects. Additionally, the response options for the hangover and withdrawal conditions were different from those options in the combined stem condition (i.e., number of times in the past year versus “yes” or “no,” respectively). This difference in conditions should be considered when interpreting the results even though the within and between-subject consistency of findings argues against this materially affecting our findings. It should be noted that although groups were not significantly different across the majority of the demographic variables, they were significantly different in terms of sex with the proportion of males being significantly higher in the withdrawal condition when compared to the

hangover condition, presumably due to chance (since subjects were randomized to the three conditions). However, given the groups have similar drinking patterns, it is unlikely that this difference significantly influenced results. To test this further, we conducted post hoc analyses to examine the effect of condition on rates of endorsement for each symptom, with sex modeled as a covariate. The interaction between sex and condition across each of the symptoms was non-significant, all p s > 0.06 (a Bonferroni adjustment would require $p < .05/8 = .00625$ to be interpreted as statistically significant), suggesting the differences in sex across conditions did not have a significant effect on findings. Future research efforts would benefit from replication of these findings across different samples, administration methods, and substances of abuse. It would also be useful to examine whether relief or avoidance are inaccurately attributed to withdrawal when they are instead used as techniques to cure hangover (e.g., drinking in the morning to alleviate hangover symptoms). Although the relief and avoidance items included in the AUDADIS are not plagued by the combined stem, there may still be misattribution that inflates estimates of withdrawal prevalence. Therefore, this would be worth exploring further.

Even with the limitations just noted, it is difficult to construe a situation where the basic findings here would not generalize, at least to some nontrivial degree to past and future research. Moreover, the ages sampled in the current sample (i.e., the college years) represent the period of life associated with the highest hazard rates for AUD diagnosis and past 12-month prevalence (Blanco et al. 2008), which suggest this is an important subsample of those diagnosed with AUD. Importantly, this is also a period highly associated with hangover (CORE Institute 2001; Prendergast 1994) likely owing to the high prevalence of binge drinking patterns (Center for Disease Control and Prevention 2012).

Future work should focus on resolving the current problem with the AUDADIS withdrawal criterion to decrease the likelihood of potential diagnostic false positives. To this end, clarifying the withdrawal stem and clearly distinguishing it from a hangover experience, similar to the stem of the SSAGA, would be highly desirable. As suggested by Karriker-Jaffe, et al. (2015), instruments used for population-based surveys may resolve problems related to measurement error if they ask about the amount of time elapsed between stopping drinking and the occurrence of the withdrawal symptoms. In addition, we believe the current experimental approach to assessing survey question validity represents a potentially valuable but underused (e.g., in comparison to cognitive interviews; Schwarz & Oyserman 2001) approach and could be applied widely to a range of assessment questions.

Additional work should also focus on strengthening the validity of the withdrawal criterion. Work by Hasin and colleagues (2000) suggest the “shakes” (i.e., tremors) symptom may have particular significance in the diagnosis of withdrawal. That is, cases in which shakes were required for diagnosis of withdrawal had stronger prognostic meaning. Test-retest data for alcoholic “shakes” has shown that the reliability of this symptom ($\kappa = .74$) is significantly better than that of other alcohol withdrawal symptoms (range, $\kappa = .53-.61$) in the DSM-IV (Hasin et al. 1997). Further, in Hasin et al. (1997), participants endorsing shakes were more likely to have many other indicators of severe alcohol dependence, suggesting the requirement of tremors for withdrawal makes withdrawal a stronger and more efficient predictor of chronicity.

Whether or not diagnosis requires tremors for a diagnosis of alcohol withdrawal has the potential to change the apparent prevalence of withdrawal as well as change the validity and/or severity of the condition that is measured. We similarly argue that consideration of the likely inflation of false positives due to the combined stem similarly could change the apparent prevalence and severity of AUD. Given the importance of withdrawal phenomena for clinical management and the central place of withdrawal in influential theories of addiction (Koob & LeMoal 2001), improving its assessment should be a high priority and the limitations of epidemiological data on its prevalence should be more fully appreciated.

Acknowledgments

The present study was supported by the NIH grants T32 AA13526 and K05AA017242 to Kenneth J. Sher.

References

- Diagnostic and Statistical Manual of Mental Disorders. 4th. American Psychiatric Association; American Psychiatric Association; Washington, DC: 2000.
- Diagnostic and Statistical Manual of Mental Disorders. 5th. American Psychiatric Association; American Psychiatric Association; Washington, DC: 2013.
- Bayard M, McIntyre J, Hill KR, Woodside J Jr. Alcohol withdrawal syndrome. *American Family Physician*. 2004; 69:1443–1450. [PubMed: 15053409]
- Becker HC. Kindling in alcohol withdrawal. *Alcohol Health and Research World*. 1998; 22:25–33. [PubMed: 15706729]
- Bierut LJ, Dinwiddie SH, Begleiter H, Crowe RR, Hesselbrock V, Nurnberger JI, Porjesz B, Schuckit MA, Reich T. Familial transmission of substance dependence: alcohol, marijuana, cocaine, and habitual smoking: a report from the Collaborative Study on the Genetics of Alcoholism. *Archives of General Psychiatry*. 1998; 55:982–988. [PubMed: 9819066]
- Blanco C, Okuda M, Wright C, Hasin DS, Grant BF, Liu SM, Olfson M. Mental health of college students and their non-college-attending peers: Results from the National Epidemiologic Study on Alcohol and Related Conditions. *Archives of General Psychiatry*. 2008; 65:1429–1437. [PubMed: 19047530]
- Blouin AG, Perez EL, Blouin JH. Computerized administration of the Diagnostic Interview Schedule. *Psychiatry Research*. 1988; 23:335–344. [PubMed: 3387505]
- Breese GR, Overstreet DH, Knapp DJ. Conceptual framework for the etiology of alcoholism: a “kindling”/stress hypothesis. *Psychopharmacology*. 2005; 178:367–380. [PubMed: 15765253]
- Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JI, Reich T, Schmidt I, Schuckit MA. A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. *Journal of Studies on Alcohol*. 1994; 55:149–158. [PubMed: 8189735]
- Caetano R, Babor TF. Diagnosis of alcohol dependence in epidemiological surveys: An epidemic of youthful alcohol dependence or a case of measurement error? *Addiction*. 2006; 101(SUPPL 1): 111–114. [PubMed: 16930167]
- Canino GJ, Bravo M, Ramirez R, Febo V, Fernandez R, Hasin DS. The Spanish Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): Reliability and concordance with clinical diagnoses in a Hispanic population. *Journal of Studies on Alcohol and Drugs*. 1999; 60:790–799.
- Casey M, Adamson G, Shevlin M, McKinney A. The role of craving in AUDs: Dimensionality and differential functioning in the DSM-5. *Drug and Alcohol Dependence*. 2012; 125:75–80. [PubMed: 22516145]
- Centers for Disease Control and Prevention (CDC). Vital signs: Binge drinking prevalence, frequency, and intensity among adults—United States, 2010. *Morbidity and Mortality Weekly Report*. 2012; 61:14. [PubMed: 22237031]

- Chatterji S, Saunders JB, Vraiti R, Grant BF, Hasin D, Mager D. Reliability of the alcohol and drug modules of the Alcohol Use Disorder and Associated Disabilities Interview Schedule – Alcohol/Drug-Revised (AUDADIS-ADR): An international comparison. *Drug and Alcohol Dependence*. 1997; 47:171–185. [PubMed: 9306043]
- Compton WM, Thomas YF, Stinson FS, Grant BF. Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry*. 2007; 64:566–576. [PubMed: 17485608]
- CORE Institute. [Accessed December 13, 2015] 2000 Statistics on Alcohol and Other Drug Use on American Campuses. 2001. Available at http://www.siu.edu/departments/coreinst/public_html/recent.html
- Embretson, SE.; Reise, SP. Item Response Theory for Psychologists. Psychology Press; New Jersey: 2000.
- Gosling SD, Vazire S, Srivastava S, John OP. Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist*. 2004; 59:93–104. [PubMed: 14992636]
- Grant, BF.; Peterson, A.; Dawson, DA.; Chou, SP. Source and Accuracy Statement for the National Longitudinal Alcohol Epidemiologic Survey (NLAES). National Institute on Alcohol Abuse and Alcoholism; Bethesda, MD: 1994.
- Grant BF, Harford TC, Dawson DA, Chou PS, Pickering RP. The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): Reliability of alcohol and drug modules in a general population sample. *Drug and Alcohol Dependence*. 1995; 39:37–44. [PubMed: 7587973]
- Grant, BF.; Dawson, DA.; Hasin, DS. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-DSM-IV Version. National Institute on Alcohol Abuse and Alcoholism; Bethesda, MD: 2001.
- Grant BF, Dawson DA, Stinson FS, Chou SP, Dufour MC, Pickering RP. The 12-month prevalence and trends in DSM-IV alcohol abuse and dependence: United States, 1991–1992 and 2001–2002. *Drug and Alcohol Dependence*. 2004; 74:223–34. [PubMed: 15194200]
- Grant BF, Dawson DA, Stinson FS, Chou SP, Kay W, Pickering RP. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): Reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug and Alcohol Dependence*. 2003; 71:7–16. [PubMed: 12821201]
- Grant, BF.; Moore, TC.; Kaplan, KD. Source and accuracy statement: Wave 1 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism; 2003.
- Grant, BF.; Kaplan, KD. Source and accuracy statement: Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). Rockville, MD: National Institute on Alcohol Abuse and Alcoholism; 2005.
- Grant, BF.; Goldstein, RB.; Chou, SP.; Saha, TD.; Ruan, WJ.; Huang, B.; Smith, SM.; Zhang, H.; Jung, J.; Pickering, RP.; Aivadyan, C.; Greenstein, E.; Hasin, DS. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition Version (AUDADIS-5). Rockville, MD: National Institute on Alcohol Abuse and Alcoholism; 2011.
- Grant, BF.; Chu, A.; Signman, R.; Amsbary, M.; Kali, J.; Sugawara, Y.; Jiao, R.; Ren, W.; Goldstein, R. [Accessed December 12, 2015] National Institute on Alcohol Abuse and Alcoholism: National Epidemiologic Survey on Alcohol and Related Conditions-III (NESARC-III): Source and accuracy statement. 2015a. Available at http://www.niaaa.nih.gov/sites/default/files/NESARC_Final_Report_FINAL_1_8_15.pdf
- Grant BF, Goldstein RB, Smith SM, Jung J, Zhang H, Chou SP, Hasin DS. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-5 (AUDADIS-5): reliability of substance use and psychiatric disorder modules in a general population sample. *Drug and Alcohol Dependence*. 2015b; 148:27–33. [PubMed: 25595052]
- Grant BF, Goldstein RB, Saha TD, Chou SP, Jung J, Zhang H, Pickering RP, Ruan WJ, Smith SM, Huang MD, Hasin DS. Epidemiology of DSM-5 alcohol use disorder: results from the National

- Epidemiologic Survey on Alcohol and Related Conditions III. *JAMA Psychiatry*. 2015c; 72:757–766. [PubMed: 26039070]
- Harford TC, Grant BF, Yi H-Y, Chen CM. Patterns of DSM-IV alcohol abuse and dependence criteria among adolescents and adults: results from the 2001 National Household Survey on Drug Abuse. *Alcoholism: Clinical and Experimental Research*. 2005; 29:810–828.
- Harford TC, Yi HY, Faden VB, Chen CM. The dimensionality of DSM-IV alcohol use disorders among adolescent and adult drinkers and symptom patterns by age, gender, and race/ethnicity. *Alcoholism: Clinical & Experimental Research*. 2009; 33:868–878.
- Hasin DS, Carpenter KM, McCloud S, Smith M, Grant BF. The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability of alcohol and drug modules in a clinical sample. *Drug and Alcohol Dependence*. 1997; 44:133–141. [PubMed: 9088785]
- Hasin DS, Greenstein E, Aivadyan C, Stohl M, Aharonovich E, Saha T, Grant BF. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-5 (AUDADIS-5): procedural validity of substance use disorders modules through clinical re-appraisal in a general population sample. *Drug and Alcohol Dependence*. 2015; 148:40–46. [PubMed: 25604321]
- Karriker-Jaffe KJ, Witbrodt J, Greenfield TK. Refining measures of alcohol problems for general population surveys. *Alcoholism: Clinical and Experimental Research*. 2015; 39:363–370.
- Keyes KM, Krueger RF, Grant BF, Hasin DS. Alcohol craving and the dimensionality of alcohol disorders. *Psychological Medicine*. 2011; 41:629–640. [PubMed: 20459881]
- Koob GF, Le Moal M. Drug addiction, dysregulation of reward, and allostasis. *Neuropsychopharmacology*. 2001; 24:97–129. [PubMed: 11120394]
- Lane SP, Steinley D, Sher KJ. Meta-analysis of DSM AUD criteria severity: Structural consistency is only “skin deep.”. *Psychological Medicine*. (in press).
- Langenbucher J, Martin CS, Labouvie E, Sanjuan PM, Bavly L, Pollock NK. Toward the DSM-V: the Withdrawal-Gate Model versus the DSM-IV in the diagnosis of alcohol abuse and dependence. *Journal of Consulting and Clinical Psychology*. 2000; 68:799–809. [PubMed: 11068966]
- Martin CS, Chung T, Langenbucher JW. How should we revise diagnostic criteria for substance use disorders in the DSM-V? *Journal of Abnormal Psychology*. 2008; 117:561–575. [PubMed: 18729609]
- Martin CS, Langenbucher JW, Chung T, Sher KJ. Truth or consequences in the diagnosis of substance use disorders. *Addiction*. 2014; 109:1773–1778. [PubMed: 24913314]
- Martin CS, Sher KJ, Chung T. Hazardous use should not be a diagnostic criterion for substance use disorders in DSM-5. *Journal of Studies on Alcohol and Drugs*. 2011; 72:685–686. [PubMed: 21683051]
- Midanik LT, Greenfield TK, Bond J. Addiction sciences and its psychometrics: The measurement of alcohol-related problems. *Addiction*. 2007; 102:1701–1710. [PubMed: 17610540]
- Miller ET, Neal DJ, Roberts LJ, Baer JS, Cressler SO, Metrik J, Marlatt GA. Test-retest reliability of alcohol measures: is there a difference between internet-based assessment and traditional methods? *Psychology of Addictive Behaviors*. 2002; 16:56–63. [PubMed: 11934087]
- Newlin DB, Pretorius MB. Sons of alcoholics report greater hangover symptoms than sons of nonalcoholics: a pilot study. *Alcohol Clinical and Experimental Research*. 1990; 14:713–716.
- Prat G, Adan A, Sanchez-Turet M. Alcohol hangover: A critical review of explanatory factors. *Human Psychopharmacology-Clinical and Experimental*. 2009; 24:259–267. [PubMed: 19347842]
- Prendergast ML. Substance use and abuse among college students: A review of recent literature. *Journal of American College Health*. 1994; 43:99–113. [PubMed: 7814772]
- Proudfoot H, Baillie AJ, Teesson M. The structure of alcohol dependence in the community. *Drug and Alcohol Dependence*. 2006; 81:21–26. [PubMed: 16005578]
- Ruan WJ, Goldstein RB, Chou SP, Smith SM, Saha TD, Pickering RP, Dawson DA, Huang B, Stinson FS, Grant BF. The alcohol use disorder and associated disabilities interview schedule-IV (AUDADIS-IV): reliability of new psychiatric diagnostic modules and risk factors in a general population sample. *Drug and Alcohol Dependence*. 2008; 92:27–36. [PubMed: 17706375]
- 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. *Psychological Medicine*. 2006; 36:931–941. [PubMed: 16563205]

- Schuckit MA, Smith TL, Daeppeen JB, Eng M, Li TK, Hesselbrock VM, Nurmberger JI, Bucholz KK. Clinical relevance of the distinction between alcohol dependence with and without a physiological component. *American Journal of Psychiatry*. 1998; 155:733–740. [PubMed: 9619144]
- Schwarz N, Oyserman D. Asking questions about behavior: Cognition, communication, and questionnaire construction. *American Journal of Evaluation*. 2001; 22:127–160.
- Skinner HA, Allen BA. Does the computer make a difference? Computerized versus face-to-face versus self-report assessment of alcohol, drug, and tobacco use. *Journal of Consulting and Clinical Psychology*. 1983; 51:267–275. [PubMed: 6841771]
- Substance Abuse and Mental Health Services Administration. Research Triangle Institute [producer], Research Triangle Park, NC. Inter-University Consortium for Political and Social Research [distributor]; Ann Arbor, MI: 2006 2006 Nov 16. U.S. Dept. of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies National Survey on Drug Use and Health, 2005 [Computer file]. ICPSR04596-v1.
- Vrasti R, Grant BF, Chatterji S, Ustun BT, Mager D, Olteanu I, Bodoi M. The reliability of the Romanian version of the alcohol module of the WHO Alcohol Use Disorder and Associated Disabilities Interview Schedule – Alcohol/Drug-Revised (AUDADIS-ADR). *European Addiction Research*. 1997; 4:144–149. [PubMed: 9852366]

Table 1

Demographic Characteristics of Sample by Condition

Characteristic	Overall (N = 497)			Hangover (n = 164)			Withdrawal (n = 167)			Combined (n = 166)		
	%	M	SD	%	M	SD	%	M	SD	%	M	SD
Age		18.7	1.2		18.7 _a	1.1		18.5 _a	1.0		18.8 _a	1.4
Sex												
Male	41.9			50.0 _a			35.5 _b			40.1 _{ab}		
Race												
White	87.1			92.1 _a			88.6 _a			86.7 _a		
African-American	7.2			7.3 _a			7.2 _a			7.2 _a		
Other _z	7.2			4.9 _a			6.6 _a			10.2 _a		
Age of first drink		15.6	1.6		15.5 _a	1.8		15.5 _a	1.4		15.8 _a	1.8
Age first drunk (n = 487)		15.8	2.8		15.8 _a	2.4		15.6 _a	3.2		16.0 _a	2.6
Past year consumption (per week)												
Drinking days/week		1.1	1.0		1.2 _a	1.0		1.1 _a	0.9		1.1 _a	1.0
Avg. drinks per occasion		5.0	2.3		5.3 _a	2.5		4.8 _a	2.1		4.9 _a	2.3
Days with 12+ drinks in single sitting		0.3	0.6		0.3 _a	0.7		0.3 _a	0.5		0.2 _a	0.6
Days "drunk"/week		0.7	0.8		0.8 _a	0.9		0.7 _a	0.7		0.7 _a	0.8

Note. Means with the same letter in their subscripts do not differ significantly from one another according to Bonferroni *t*-tests of differences between means for continuous variables and chi square tests for categorical variables.

_z Group designation "Other" includes those who identified as being in more than one racial category (n = 10) therefore the numbers do not sum to 100.

Chi Square and Logistic Regression Comparisons for AUDADIS Withdrawal Symptoms Across Conditions

Table 2

Symptom	χ^2 (2)	Hangover (% Endorsed; <i>n</i> = 164)	Withdrawal (% Endorsed; <i>n</i> = 167)	Combined (% Endorsed; <i>n</i> = 166)	Odds Ratio for Withdrawal vs Combined	Odds Ratio for Hangover vs Combined
Sleep Trouble	40.0 *	27.8 _a	3.7 _b	28.1 _a	0.1 *	1.0
Nausea/Vomiting	131.5 *	61.1 _a	6.2 _b	60.5 _a	0.0 *	1.0
Restless	47.1 *	34.0 _a	3.7 _b	24.0 _a	0.1 *	1.6
Sweat/Heart Beat Fast	33.0 *	23.5 _a	2.5 _b	22.2 _a	0.1 *	1.1
Anxious	21.0 *	11.8 _a	4.3 _a	21.0 _b	0.2 *	0.5
Shakes	12.7 *	14.8 _a	3.7 _b	13.8 _a	0.2 *	1.1
Delirium Tremens	14.2 *	11.7 _a	1.2 _b	7.8 _a	0.1	1.6
Seizures	10.7 *	8.0 _a	0.6 _a	7.2 _a	0.1	1.1
Headache	111.3 *	49.7 _a	8.0 _b	62.9 _a	0.1 *	0.6
Average Number of Symptoms Endorsed (SD)		3.2 (3.0)	0.5 (1.6)	3.3 (2.7)		

Note. Cells sharing the same subscript do not differ significantly from one another at the $p < .0056$ level. Although the AUDADIS-IV assesses headache, it is not one of the 8 symptoms used by the instrument to diagnose alcohol withdrawal.

* $p < .0056$ (setwise alpha)

Table 3

Withdrawal Symptoms Within the Combined Stem Condition, n = 166 (Prevalence of Hangover and Withdrawal Among Those Who Endorse the Combined Stem)

Symptom	Positive Endorsement (%)	Hangover (% Endorsed)	Withdrawal (% Endorsed)	Ratio	p-value
Sleep Trouble	28.1	33.4 (92.9)	0.4 (14.3)	6.5:1	<0.01
Nausea/Vomiting	60.5	82.3 (96.9)	0.3 (11.5)	247.3:1	<0.01
Restless	24.0	23.9 (96.9)	0.9 (15.6)	26.6:1	<0.01
Anxious	21.0	15.7 (82.8)	1.7 (34.5)	9.2:1	<0.01
Sweat/Heart Beat Fast	22.2	27.1 (96.8)	0.1 (9.7)	271.0:1	<0.01
Shakes	13.8	12.9 (85.7)	0.9 (28.6)	14.3:1	<0.01
Delirium Tremens	7.8	3.6 (60.0)	1.6 (4.0)	2.3:1	0.50
Seizures	7.2	1.6 (54.5)	3.6 (72.7)	0.4:1	0.50
Headache	62.9	89.3 (97.1)	0.3 (9.8)	297.7:1	<0.01

Note. p-value is based on prevalence of hangover and withdrawal based on McNemar's Test. Rates of endorsement for hangover and withdrawal are based on the overall subsample (n = 166). Numbers in parentheses indicate the prevalence of the symptom conditional upon positive endorsement of the combined stem. Ratio is the ratio of the hangover endorsement rate to the withdrawal endorsement rate. In some cases the sum of "hangover" endorsements and "withdrawal" endorsements is less than the "combined" endorsement because a participant might have endorsed the combined stem but when probed, did not endorse either option. Although the AUDADIS-IV assesses headache, it is not one of the 8 symptoms used by the instrument to diagnose alcohol withdrawal.