Published in final edited form as:

Drug Alcohol Depend. 2014 August 1; 141: 39–43. doi:10.1016/j.drugalcdep.2014.05.004.

A Test of the DSM-5 Severity Scale for Alcohol Use Disorder

Tera L. Fazzino^{1,2,*}, Gail L. Rose², Keith B. Burt¹, and John E. Helzer²

¹Department of Psychology, University of Vermont, 2 Colchester Avenue, Burlington, VT 05401, USA

²Department of Psychiatry, University of Vermont, 1 South Prospect Street, Burlington, VT 05401, USA

Abstract

BACKGROUND—For the DSM-5-defined alcohol use disorder (AUD) diagnosis, a tricategorized scale that designates mild, moderate, and severe AUD was selected over a fully dimensional scale to represent AUD severity. The purpose of this study was to test whether the DSM-5-defined AUD severity measure was as proficient a predictor of alcohol use following a brief intervention, compared to a fully dimensional scale.

METHODS—Heavy drinking primary care patients ($N=246$) received a physician-delivered brief intervention (BI), and then reported daily alcohol consumption for six months using an Interactive Voice Response (IVR) system. The dimensional AUD measure we constructed was a summation of all AUD criteria met at baseline (mean = 6.5; $SD = 2.5$). A multi-model inference technique was used to determine whether the DSM-5 tri-categorized severity measure or a dimensional approach would provide a more precise prediction of change in weekly alcohol consumption following a BI.

RESULTS—The Akaike information criterion (AIC) for the dimensional AUD model ($AIC=7623.88$) was four points lower than the tri-categorized model ($AIC=7627.88$) and weight of evidence calculations indicated there was 88% likelihood the dimensional model was the better approximating model. The dimensional model significantly predicted change in alcohol consumption ($p=.04$) whereas the DSM-5 tri-categorized model did not.

© 2014 Elsevier Ireland Ltd. All rights reserved.

*Corresponding Author: Tera L. Fazzino, Health Behavior Research Center, Department of Psychiatry, University of Vermont, UHC Campus 4570H3, 1 South Prospect Street, Burlington, VT 05401 USA, tfazzino@uvm.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Contributors

Authors Dr. Rose and Dr. Helzer designed and conducted the parent study from which the data for this secondary analysis were derived. Ms. Fazzino managed the literature searches and summaries of previous related work. Ms. Fazzino and Dr. Burt undertook the statistical analysis. Ms. Fazzino wrote the first draft of the manuscript and all authors contributed to revisions and the final version of the manuscript. All authors have approved the final manuscript.

Conflicts of Interest

The authors have no conflicts of interest to declare.

CONCLUSION—A dimensional AUD measure was superior, detecting treatment effects that were not apparent with tri-categorized severity model as defined by the DSM-5. We recommend using a dimensional measure for determining AUD severity.

Keywords

DSM-5; diagnosis; alcohol use disorder severity; dimensional; alcohol use disorder; multi-model inference

1. INTRODUCTION

With the publication of the DSM-5, several evidence-based changes to the alcohol use disorder (AUD) diagnosis have been made. Based on recommendations from researchers, DSM-IV-defined Alcohol Abuse (AA) and Alcohol Dependence (AD) were combined into a single AUD diagnosis, the legal problems criterion was removed, and a craving criterion was added (Hasin et al., 2013).

One change that was widely recommended (Tarter et al., 1992; Muthén, 2006; Helzer et al., 2006a, 2006b, 2007) but not incorporated into the DSM-5 was a fully dimensional diagnostic AUD scale. Such a scale would allow for severity to be measured based on a count of the number of symptoms endorsed. Instead, a tri-categorized severity scale was used that distinguishes between mild, moderate, and severe dependence (American Psychiatric Association, 2013). Two to three AUD symptoms are considered as mild dependence, 4–5 symptoms define moderate dependence, and 6 or more symptoms are considered severe dependence (American Psychiatric Association, 2013). In addition, while AA and AD were combined to form a single AUD diagnosis, the DSM-5 retains the categorical dichotomous diagnostic structure based on the presence of two or more symptoms. Thus, a categorical diagnostic structure has been retained for both AUD diagnosis as well as the AUD severity measure.

In the original theory of alcohol dependence proposed by Edwards and Gross (1976) the disorder was conceptualized as a dimensional construct. Specifically, Edwards and Gross proposed that alcohol dependence exists on a continuum of severity and emphasized the importance of recognizing alcohol dependence in degrees in both research and clinical work. A dimensional measure that counts the number of criteria an individual meets provides information about diversity in symptom presentation and is valuable for both clinical and research purposes (Hasin et al., 2006; Helzer et al., 2006b).

A dimensional scale is also preferable to a categorical measure because it can substantially increase statistical power; the high statistical costs of categorizing a quantitative variable have been demonstrated (Cohen, 1983; MacCallum et al., 2002). In this regard, we have previously shown that a dimensional diagnostic measure is superior to a DSM-IV defined dichotomous alcohol dependence diagnosis for predicting change in alcohol consumption following a brief intervention (BI; Fazzino et al., 2014.). Dimensional measurement also better accommodates instances of diagnostic comorbidity. With dimensional scales, comorbidity between variables can be discussed in terms of their degree of correlation, compared to with categorical designations which would require discussion of individuals

meeting criteria for more than one disorder (Goldberg, 2000). Finally, as demonstrated in an international study with clinicians from 64 countries, dimensional constructs of mental disorders better reflected clinicians' conceptualizations of mental disorders than did categorical conceptualizations (Roberts et al., 2012).

While the severity measure was included in the DSM-5 to distinguish between levels of AUD (Hasin et al., 2013), neither the DSM-5 itself nor the article detailing rationale for DSM-5 changes (Hasin et al., 2013) indicates how the severity level cut-off points were derived. Borges et al. (2011) analyzed potential AUD severity cut off points for the DSM-5 using Item Response Theory and suggested using a dual category severity scale with 2–3 AUD symptoms designating moderate AUD and 4–11 indicating severe AUD. However, the tri-categorized severity scale that was chosen for the DSM-5 does not map directly onto the Borges et al. findings. Despite rigorous statistical evaluation of other changes incorporated in DSM-5 AUD diagnosis, it appears that the specific severity cut points as delineated in the DSM-5 may not have been thoroughly statistically evaluated. Historically, a substantial criticism of DSM-defined categorical disorders is that diagnostic criteria and thresholds have been based on clinical observation or historical convention, not statistical evaluation (Goldberg, 2000; Helzer et al., 2006b; Jones, 2012; Kessler, 2002; Widiger and Simonsen, 2005) which has led to statistical analysis of DSM-defined constructs post-hoc. In this regard, it appears that the DSM-5 severity scale represents a continuation of this problem.

It is currently unclear whether the tri-categorized severity measure is as proficient a predictor of alcohol use treatment outcomes as a fully dimensional scale. Due to the categorical nature of the severity scale, we expected that the severity measure would not adequately consider variability in the data that would be necessary to predict alcohol use as precisely as a fully dimensional scale.

The purpose of the current study was to compare goodness of fit of the DSM-5 severity measure to a full dimensional AUD scale for predicting alcohol use following a physician-delivered BI. Specifically, we compared goodness of fit of the tri-categorized severity measure to a full dimensional measure of AUD for predicting first week alcohol consumption following a BI (intercept) and change in weekly alcohol consumption (slope) over four weeks following a BI.

2. METHODS

Data used in the current manuscript were obtained from a study that evaluated 6 months of self-monitoring via Interactive Voice Response (IVR) following a BI for alcohol use in a primary care setting (Helzer et al., 2008). The main objective of the original study was to determine if self-monitoring via IVR with or without monthly feedback about alcohol use would produce a greater reduction in alcohol consumption compared to no self-monitoring following a BI. Feedback was provided at the end of each month.

2.1 Participants

Participants were recruited from 15 primary care offices in the Burlington, Vermont metropolitan area. Primary care providers conducted brief interventions with heavy drinking

primary care patients and referred patients who were willing to participate in the randomized trial to the research staff. Participants were included in the study if they reported recent (past 3 month) alcohol consumption beyond the NIAAA guidelines for low risk drinking: 1) average daily or weekly alcohol use exceeding 2 drinks per day/14 per week for men or 1 per day/7 per week for women, or 2) 5 or more drinks in one day for men or 4 or more for women (National Institute of Alcohol and Alcohol Abuse, 2005). Both dependent and not dependent individuals as defined by DSM-IV categorical criteria were included in the sample. Exclusion criteria were current (past year) DSM-IV diagnosis of substance dependence other than alcohol, nicotine, or marijuana; a current diagnosis of psychosis; or a recent initiation or change in antidepressant medication that could affect alcohol consumption.

2.2 Procedure

Research personnel first contacted each patient by telephone to describe the study and then scheduled patients interested in participating in the study for an in-person informed consent and assessment at our research office. After a complete description of the study to the subjects, written informed consent was obtained. Detailed study procedures and the full assessment battery were presented previously (Helzer et al., 2008). Participants received a training session during which they were instructed on reporting standard drink volumes and oriented to using the IVR. Participants were provided a toll-free, 24 hour access phone number to contact the IVR and were asked to call daily for 6 months. The IVR call was a 2-minute questionnaire that assessed alcohol consumption (number of standard servings of beer, liquor, and wine assessed separately), craving intensity, reasons for drinking/abstaining from drinking, questions about psychological status (stress, happiness, anger, sadness), physical health, relationship with partner, partner alcohol use, and whether the participant was intoxicated at the time of the call. All IVR questions inquired about the previous 24 hour period (i.e., “yesterday”) to ensure a consistent reporting period.

2.3 Predictor Variables

Symptoms of AUD were assessed with the Composite International Diagnostic Interview-Substance Abuse Module for DSM-IV (CIDI-SAM; Cottler et al., 1989). The DSM-5 dimensional AUD diagnosis was determined by counting the number of AUD criteria (0–11) met in the past 12 months. DSM-5 defined AUD symptoms included: 1) tolerance, 2) withdrawal, 3) substance taken in larger amounts/longer period than intended, 4) persistent desire or unsuccessful attempts to decrease/control use, 5) great deal of time spent obtaining, using or recovering from effects of alcohol, 6) social, occupational, or recreational activities given up or reduced because of use, 7) use despite knowledge of physical or psychological problems caused or exacerbated by use, 8) recurrent failure to fulfill major role obligations, 9) recurrent use in hazardous situations, 10) craving/strong desire to use the substance, 11) continued use despite social/interpersonal problems. A dimensional measure using a simple symptom count was chosen over a dimensional scale with weighted criteria because Dawson et al. (2010) previously demonstrated that these two measures are similarly proficient in predicting alcohol use.

The craving criterion in the dimensional and severity scales was approximated because the CIDI-SAM version used in the study did not assess craving. Craving in the IVR was measured using the following prompt: “Rate your urge to drink yesterday on a scale of 0 to 9, with 0 being no urge to drink and 9 being the strongest urge ever to drink.” We determined the presence of craving from the first IVR report that participants made the day after they completed their diagnostic interview. We reasoned that if a participant reported craving in the first IVR report, they likely would have reported craving during the diagnostic interview the previous day. In order to assess what score threshold (1–9) to consider positive we used the mean daily craving rating for the sample. Mean craving rating across the 30 days of the study was 3.76 (SD = 2.31). Thus we defined a positive report of craving to be a score of 3 or higher on that first IVR call. In DSM-5 the actual “craving criterion” reads as follows: “Craving, or a strong desire or urge to use alcohol.” Therefore, we also operationalized the score threshold as any non-zero craving report in the first IVR call since that seemed to most closely match the actual DSM-5 definition. We used these thresholds to create corresponding dichotomous variables (craving/no craving). Below we report analyses using both definitions.

The DSM-5 AUD severity measure was determined based on DSM-5 specified severity criteria of 2–3 positive symptoms for mild dependence, 4–5 symptoms for moderate dependence, and 6 or more for severe (American Psychiatric Association, 2013).

2.4 Outcome Variable

Alcohol use was assessed daily in the IVR survey with the following prompt, using a separate question for each type of alcohol: “How many [beers/drinks containing liquor/glasses of wine] did you have yesterday?” Validity of previous day alcohol consumption reported via the IVR has previously been demonstrated (Searles et al., 1995). The total weekly alcoholic drinks variable was computed by summing the total number of alcoholic drinks reported to the IVR in each 7 day period. First week total alcohol consumption was operationalized as the total number of alcoholic drinks consumed in the first study week (model intercept), and change in alcohol consumption was modeled as the change in total drinks per week (model slope).

While participants in the parent study reported alcohol use for 6 months, analyses for the current study were restricted to the first month because the parent study observed the strongest treatment effects in the first month following the BI (Helzer et al., 2008), consistent with other BI efficacy studies in the literature (Moyer et al., 2002). Restricting our time points to four also allowed us to evaluate a timeframe that was appropriate for the data analytic technique used (please see Data Analysis section).

2.5 Data Analysis

A multi-model inference technique was used for all analyses because this technique provides a comprehensive evaluation of the likelihood of superior model fit with the data. Standard null-hypothesis significance testing was not the primary method of analysis, because binary significant/not significant estimates use an arbitrary .05 cut point to determine significance and do not evaluate model fit (Cowles and Davis, 1982). However, standard model

coefficients are provided to consider whether the predictor variables were meaningful predictors of alcohol consumption beyond chance.

2.6 Linear mixed models description

Four different models with AUD measure (severity or dimensional) predicting total drinks per week were estimated using linear modeling with generalized least squares (GLS) to account for correlated observations. Analyses were conducted in R (R Development Core Team, 2012) with the *gls()* function from the *nlme* package (Pinheiro et al., 2012) utilizing a longitudinal data environment which allows clustering by participant identification number and sorting by study week from 1 through 4.

We evaluated exchangeable, unstructured, and autoregressive correlation structures in the data and determined that an autoregressive correlation structure fit the data best based on Akaike's Information Criteria (AIC; Akaike, 1973, 1974) and weight of evidence estimates. Thus, all models assumed an autoregressive correlation structure.

Two pairs of GLS models were estimated to evaluate whether the DSM-5-defined AUD severity measure or a dimensional AUD measure was the best fit to the data in predicting first week alcohol consumption (intercept) and change in alcohol consumption per week over four weeks (slope). Models were constructed with AUD measure predicting total drinks consumed the first study week and change in weekly total drinks consumed.

The two pairs of GLS models differed in the way the craving criterion was operationalized for the AUD measure. In the first set of models, both the severity model A1 and dimensional model A2 operationalized the craving criterion as a craving rating of 3 or higher. In the second set of models, severity model A3 and dimensional model A4 both included a craving criterion conceptualized as any non-zero craving rating.

In addition to the base models, we reran our analyses with age and gender included as covariates due to their theoretical and statistical relevance to our models (Grant, 1997; Nolen-Hoeksema, 2004).

Four participants who reported very high levels of weekly alcohol use (80–155 drinks per week) skewed our sample (skewness=1.34). Therefore, we also reran our analyses excluding these outliers to determine whether they unduly influenced the results of our models.

2.7 Multi-model inference technique description

Multi-model inference technique using maximum likelihood estimation (MLE) was used to compare model fit to the data. MLE uses a likelihood function to determine the rank ordered likelihood that each model is the best fitting model based on the observed data (Long, 2011). The better fitting model using MLE was determined with two sets of criteria: 1) comparing the AIC estimates for each pair of models (A1 compared with A2, A3 compared with A4); and 2) comparing weight of evidence estimates of the two pairs of models. A lower AIC is considered a better fit to the data (Long, 2011) and the magnitude of the difference between the AIC values of each model, or delta, indicates the degree of support for the preferred model. While no precise cutoff points have been specified for the delta, differences between

0–2 are generally considered minimal, differences 4–7 are generally considered moderate to strong, and differences of 8–10 or above are generally considered very strong (Burnham and Anderson, 2002). Weight of evidence estimates indicate the probability that a model is the best approximating model to the data and provides this estimate for each model out of 100% (Long, 2011). For all analyses, missing data were treated as missing at random and MLE methods were employed to make use of all available data.

3. RESULTS

3.1 Participant Information

The final sample included 246 participants (166 men, 80 women) with mean age of 46 years ($SD = 13$, range = 21–82). Ninety-seven percent of participants were Caucasian/non-Hispanic, 76% reported being employed full time, and the sample had completed a mean of 15 years of education.

3.2 Descriptive Statistics

The percentage of participants who met criteria for DSM-5 AUD severity is presented in Table 1. The mean number of AUD symptoms met for dimensional model A2 was 6.5 ($SD = 2.5$) and the number of positive dependence criteria was normally distributed. The mean number of AUD symptoms met for dimensional model A4 was 6.7 ($SD = 2.4$). Participants drank on 94% of the four study weeks and the mean number of drinks per week across the study was 26.5 ($SD = 5.3$, range = 0–155).

3.3 IVR Compliance

Of the 246 participants in the sample, 70% completed 100% of their scheduled calls. Fourteen percent of participants missed a small percentage of scheduled calls (between 10–30%), 6% missed greater than 50% of scheduled calls, and 10% made 0 calls. Collectively, participants completed a total of 6,188 IVR reports for the four weeks.

3.4 Multi-model inference

Based on Burnham and Anderson's (2002) criteria for delta AIC values there was moderate to strong evidence that dimensional model A2 was a better fit for predicting alcohol use compared to its respective DSM-5 severity model (A1; Table 2). The weight of evidence estimate (Long, 2011) for the A2 dimensional model was 88%, also indicating that the dimensional model was the better fitting compared to its respective severity model ($A2 = 0.88$). The weight of evidence estimate for the severity model was low (12%), indicating low support for the severity model being the better fit ($A1 = 0.12$). Results for the second pair of models indicated even stronger support for the dimensional model being the better fit than the severity model (94% and 6% respectively; see Table 2).

The magnitude of the AIC delta between the DSM-5 severity models and their corresponding dimensional models increased with the addition of covariates age and gender and with the exclusion of outliers (Model Set 1: delta AIC = 5.22, Model Set 2: delta AIC = 6.18), indicating even stronger support for the dimensional models being better fits to the data than their corresponding severity models. With these parameters, the weight of

evidence for the dimensional models slightly increased to 93% for dimensional model A2 ($A2=.93$) and 96% for dimensional model A4 ($A4=0.96$).

3.5 Model Coefficients

The dimensional models (A2, A4) and severity model A1 significantly predicted first week alcohol use; however, the relationship was stronger in the dimensional models (Table 3). Severity model A3 did not significantly predict first week alcohol use. The dimensional A2 model-predicted number of drinks consumed the first week was 20.37 for participants with 0 symptoms of AUD, with 1.61 additional drinks for each additional symptom endorsed.

Both dimensional models significantly predicted change in weekly alcohol consumption over the four weeks following the brief intervention, whereas the severity models did not (Table 3). The A2 dimensional model-predicted change was 0.92 fewer drinks per week for participants with 0 symptoms, with an additional 0.29 drink decrease per week per additional symptom. Thus, the decrease in alcohol consumption was more prominent for participants with higher dimensional scores compared to those with lower dimensional scores. Results were similar for the dimensional A4 model (Table 3).

Adding covariates age and gender to the models and excluding outliers did not change the significance of any model coefficients and did not substantially change the magnitude of the effects.

4. DISCUSSION

The purpose of the current study was to test whether the tri-categorized measure of AUD severity as specified by the DSM-5, was as proficient a predictor of 30 day alcohol consumption following a brief intervention compared to a fully dimensional measure of AUD severity. While there were several changes made to the DSM-5 AUD criteria that were informed by research, a tri-categorized measure of AUD severity was selected over a fully dimensional AUD severity measure that was recommended by researchers (Tarter et al., 1992; Muthén, 2006; Helzer et al., 2006b, 2006a, 2007). As hypothesized, the fully dimensional model was a substantially better fit for predicting alcohol use following a BI compared to the DSM-5 tri-categorized severity measure. The dimensional scale detected significant treatment effects following the BI, whereas the severity measure did not. It is likely that the dimensional measure was a better predictor of change in alcohol use because it considered variability in symptom severity that was not afforded by the categorical severity scale (Cohen, 1983; MacCallum et al., 2002).

It is not a foregone conclusion that the dimensional model would provide a better fit for the data simply because the dimensional scale contains a larger range of values than does the categorical severity measure. If the variability within each criterion item of the dimensional scale did not have a substantial relationship with alcohol consumption, the tri-categorized severity model would be the more parsimonious and would have been the better fitting model. AIC penalizes for estimation of additional model parameters; for models that are equivalent representations of data, the more parsimonious model will obtain a lower AIC value indicating a better model fit.

It appears that the selection of the tri-categorized severity scale over a fully dimensional measure represents a significant missed opportunity for the field and may hamper the detection of treatment effects in other studies, as it did in the current study.

A limitation to our study was that we approximated the DSM-5 craving criterion because the DSM-IV diagnostic interview conducted at the time of data collection did not evaluate craving. We recognize that our single day measure of craving is an imperfect substitution for the DSM-5 craving criterion. For this reason, we evaluated two variations of our craving criterion (one conservative, one less conservative) and results were very similar in both models. Another limitation is that a small percentage (10%) of the total sample met DSM-5 criteria for mild AUD. However, the criteria for mild AUD (2–3 symptoms) encompass a small range on the AUD scale. In addition, 30 days of data were collected from each participant and our overall sample size was large. Thus, even with the small percentage of participants with mild AUD, the sample was sufficient for determining the predictive properties of the severity scale.

The current study had several strengths, including a large sample size and a nearly complete set of daily process data on alcohol consumption. In addition, the dimensional symptoms were normally distributed in our sample, which allowed our models to consider full range of variability in diagnostic severity. Also, we collected self-report data via IVR. Kobak et al. (1997) demonstrated that respondents are more willing to disclose sensitive information about alcohol consumption to an IVR compared to a live interviewer. We did use MLE to account for the small amount of missing data, but it was not necessary to use data imputation because our call compliance was high.

Our results should be tested in other samples including a random, nationally representative sample of the general population and a sample that contains a higher percentage of individuals with lower levels of AUDs. However, based on our results it is reasonable to suggest that researchers use a fully dimensional diagnostic approach when predicting treatment outcomes. Researchers may observe poorer results when using the DSM-5 severity scale to predict change in alcohol consumption compared to a fully dimensional measure of AUD severity. We recognize that using a categorized AUD severity scale may be useful, or even preferred in some cases. However, if it were desirable to have a categorized severity scale, it should be derived statistically, tested in multiple datasets, and should exhibit predictive ability for relevant alcohol treatment outcomes. For these reasons, we recommend using a dimensional measure based on a count of positive symptoms for determining AUD severity.

Acknowledgments

Role of Funding Source

The study was funded by the National Institute on Alcohol Abuse and Alcoholism (R01-AA11954, Dr. Helzer, PI). NIAAA grant R01AA018658 supported the following researchers' time during manuscript preparation: Ms. Fazzino, Dr. Rose, and Dr. Helzer. Dr. Burt's time was supported by Psychology Department funds at the University of Vermont. The NIAAA had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

REFERENCES

- Akaike, H. Information theory and an extension of the maximum likelihood principle. In: Parzen, E.; Tanabe, K.; Kitagawa, G., editors. *Selected Papers of Hirotugu Akaike*. New York, New York, NY: Springer; 1973. p. 199-213.
- Akaike H. A new look at the statistical model identification. *IEEE Trans. Autom. Control*. 1974; 19:716–723.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5)* [WWW Document]. Barnes Noble. 2013 URL <http://www.barnesandnoble.com/w/diagnostic-and-statistical-manual-of-mental-disorders-5th-edition-american-psychiatric-association/1114266954>.
- Borges G, Cherpitel CJ, Ye Y, Bond J, Cremonte M, Moskalewicz J, Swiatkiewicz G. Threshold and optimal cut-points for alcohol use disorders among patients in the emergency department. *Alcohol. Clin. Exp. Res*. 2011; 35:1270–1276. [PubMed: 21352249]
- Burnham, KP.; Anderson, DR. *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*. 2nd ed.. New York: Springer; 2002.
- Cohen J. The cost of dichotomization. *Appl. Psychol. Meas*. 1983; 7:249–253.
- Cottler LB, Robins LN, Helzer JE. The reliability of the CIDI-SAM: a comprehensive substance abuse interview. *Br. J. Addict*. 1989; 84:801–814. [PubMed: 2758153]
- Cowles M, Davis C. On the origins of the .05 level of statistical significance. *Am. Psychol*. 1982; 37:553–558.
- Dawson DA, Saha TD, Grant BF. A multidimensional assessment of the validity and utility of alcohol use disorder severity as determined by item response theory models. *Drug Alcohol Depend*. 2010; 107:31–38. [PubMed: 19782481]
- Edwards G, Gross MM. Alcohol dependence: provisional description of a clinical syndrome. *Br. Med. J*. 1976; 1:1058–1061. [PubMed: 773501]
- Fazzino TL, Rose GL, Burt KB, Helzer JE. Comparison of categorical alcohol dependence versus a dimensional measure for predicting weekly alcohol use in heavy drinkers. *Drug Alcohol Depend*. 2014; 136:121–126. [PubMed: 24485060]
- Goldberg D. Plato versus Aristotle: categorical and dimensional models for common mental disorders. *Compr. Psychiatry*. 2000; 41:8–13. [PubMed: 10746898]
- Grant BF. Prevalence and correlates of alcohol use and DSM-IV alcohol dependence in the United States: results of the National Longitudinal Alcohol Epidemiologic Survey. *J. Stud. Alcohol Drugs*. 1997; 58:464.
- Hasin DS, Liu X, Alderson D, Grant BF. DSM-IV alcohol dependence: a categorical or dimensional phenotype? *Psychol. Med*. 2006; 36:1695–1705. [PubMed: 17038207]
- Hasin DS, O'Brien CP, Auriacombe M, Borges G, Bucholz K, Budney A, Compton WM, Crowley T, Ling W, Petry NM, Schuckit M, Grant BF. DSM-5 criteria for substance use disorders: recommendations and rationale. *Am. J. Psychiatry*. 2013; 170:834–851. [PubMed: 23903334]
- Helzer JE, Bucholz KK, Gossop M. A dimensional option for the diagnosis of substance dependence in DSM-V. *Int. J. Methods Psychiatr. Res*. 2007; 16(Suppl. 1):S24–S33. [PubMed: 17623392]
- Helzer JE, Kraemer HC, Krueger RF. The feasibility and need for dimensional psychiatric diagnoses. *Psychol. Med*. 2006a; 36:1671–1680. [PubMed: 16907995]
- Helzer JE, Rose GL, Badger GJ, Searles JS, Thomas CS, Lindberg SA, Guth S. Using interactive voice response to enhance brief alcohol intervention in primary care settings. *J. Stud. Alcohol Drugs*. 2008; 69:251–258. [PubMed: 18299766]
- Helzer JE, van den Brink W, Guth SE. Should there be both categorical and dimensional criteria for the substance use disorders in DSM-V? *Addict. Abingdon Engl*. 2006b; 101(Suppl. 1):17–22.
- Jones KD. Dimensional and cross-cutting assessment in the DSM-5. *J. Couns. Dev*. 2012; 90:481–487.
- Kessler RC. The categorical versus dimensional assessment controversy in the sociology of mental illness. *J. Health Soc. Behav*. 2002; 43:171–188. [PubMed: 12096698]

- Kobak KA, Taylor LvH, Dottl SL, Greist JH, Jefferson JW, Burroughs D, Katzelnick DJ, Mandell M. Computerized screening for psychiatric disorders in an outpatient community mental health clinic. *Psychiatr. Serv.* 1997; 48:1048–1057. [PubMed: 9255838]
- Long, JD. *Longitudinal Data Analysis for the Behavioral Sciences Using R*. Thousand Oaks: SAGE Publications, Inc; 2011.
- MacCallum RC, Zhang S, Preacher KJ, Rucker DD. On the practice of dichotomization of quantitative variables. *Psychol. Methods.* 2002; 7:19–40. [PubMed: 11928888]
- Moyer A, Finney JW, Swearingen CE, Vergun P. Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. *Addiction.* 2002; 97:279–292. [PubMed: 11964101]
- Muthén B. Should substance use disorders be considered as categorical or dimensional? *Addict. Abingdon Engl.* 2006; 101(Suppl. 1):6–16.
- National Institute of Alcohol and Alcohol Abuse. *Helping Patients With Alcohol Problems: A Clinician's Guide*. Bethesda, MD: NIH; 2005. (No. 05-3769)
- Nolen-Hoeksema S. Gender differences in risk factors and consequences for alcohol use and problems. *Clin. Psychol. Rev.* 2004; 24:981–1010. [PubMed: 15533281]
- Pinheiro J, Bates D, DebRoy S, Sarkar D. R Development Core Team. *nlme: Linear and Nonlinear Mixed Effects Models*. 2012
- R Development Core Team. *R: A language and environment for statistical computing*. R package version 3.1-103. 2012
- Roberts MC, Reed GM, Medina-Mora ME, Keeley JW, Sharan P, Johnson DK, Mari J, de J, Ayuso-Mateos JL, Gureje O, Xiao Z, Maruta T, Khoury B, Robles R, Saxena S. A global clinicians' map of mental disorders to improve ICD-11: analysing meta-structure to enhance clinical utility. *Int. Rev. Psychiatry Abingdon Engl.* 2012; 24:578–590.
- Searles JS, Perrine MW, Mundt JC, Helzer JE. Self-Report of Drinking Using Touch-Tone Telephone: Extending the Limits of Reliable Daily Contact. *J. Stud. Alcohol Drugs.* 1995; 56:375.
- Tarter RE, Moss HB, Arria A, Mezzich AC, Vanyukov MM. The psychiatric diagnosis of alcoholism: critique and proposed reformulation. *Alcohol. Clin. Exp. Res.* 1992; 16:106–116. [PubMed: 1558290]
- Widiger TA, Simonsen E. Alternative dimensional models of personality disorder: finding a common ground. *J. Personal. Disord.* 2005; 19:110–130.

Table 1

Percent of participants who met DSM-5 defined mild, moderate and severe Alcohol Use Disorder (AUD).

Model	Mild	Moderate	Severe
DSM-5 AUD Severity Model (A1)	11%	22%	65%
DSM-5 AUD Severity Model (A3)	10%	19%	69%

Note. Model A1 operationalized the AUD craving criterion as craving ratings of 3 or higher and Model A3 included the craving criterion based on ratings of 1 or higher. For each model, 2% of participants did not meet criteria for an AUD.

Table 2

Model fit based on AIC values and weight of evidence estimates from multi-model inference analysis.

Model Comparison 1	AIC Value	Delta AIC	Weight of Evidence
DSM-5 AUD Severity Model (A1)	7627.88	-	12%
DSM-5 Dimensional AUD Model (A2)	7623.88	4.00	88%
Model Comparison 2			
DSM-5 AUD Severity Model (A3)	7630.13	-	6%
DSM-5 Dimensional AUD Model (A4)	7624.74	5.39	94%

Note. AIC: Akaike's Information Criterion. Lower values suggest *better* model fit; Delta AIC: number of AIC points difference between the denoted dimensional and severity models. Weight of evidence estimate is based out of 100% for each pair of models; AUD: alcohol use disorder. Models in comparison 1 (A1, A2) included the craving criterion based on ratings of 3 or higher. Models in comparison 2 (A3, A4) included the craving criterion based on ratings of 1 or higher.

Table 3

DSM-5 specified AUD severity measure and a dimensional AUD measure predicting change in weekly alcohol consumption during the first four weeks following a brief intervention.

DV: Total weekly alcoholic drinks consumed				
	<i>b</i>	SE	<i>p</i>	95% CI (low, high)
Severity AUD Model A1				
Intercept	21.32	4.66	0.01	12.16, 30.47
Diagnosis	3.87	1.79	0.03	0.35, 7.38
Week	0.44	1.14	0.70	−1.80, 2.68
Diagnosis*week	−0.56	0.44	0.20	−1.42, 0.30
Severity AUD Model A3				
Intercept	23.58	4.97	0.01	13.84, 33.33
Diagnosis	2.89	1.87	0.12	−0.79, 6.56
Week	0.23	1.21	0.85	−2.14, 2.60
Diagnosis*week	−0.47	0.46	0.31	−1.36, 0.43
Dimensional AUD Model A2				
Intercept	20.37	3.98	0.01	12.57, 28.18
Diagnosis	1.61	0.57	0.01	0.49, 2.72
Week	0.92	0.97	0.35	−0.99, 2.82
Diagnosis*week	−0.29	0.14	0.04	−0.56, −0.01
Dimensional AUD Model A4				
Intercept	20.66	4.09	0.01	12.63, 28.69
Diagnosis	1.53	0.57	0.01	0.41, 2.66
Week	0.91	1.00	0.36	−1.04, 2.87
Diagnosis*week	−0.28	0.14	0.05	−0.55, −0.004

Note. DV: dependent variable; *b* regression coefficient; SE: standard error; *p* *p*-value; CI: confidence interval; AUD: alcohol use disorder.