NIH Public Access
Author Manuscript
Drug Alcohol Depend. Author manuscript; available in PMC 2008 November 2.

Published in final edited form as:
Drug Alcohol Depend. 2007 November 2; 91(1): 85–90.

Reliability of DSM-IV Diagnostic Criteria Using the Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA)

Amira Pierucci-Lagha[a], **Joel Gelernter**[b], **Grace Chan**[a], **Albert Arias**[a], **Joseph F. Cubells**[c], **Lindsay Farrer**[d], and **Henry R. Kranzler**[a,*]

a*Alcohol Research Center, Department of Psychiatry, University of Connecticut School of Medicine, 263 Farmignton Avenue, Farmington Connecticut CT 06030-2103*

b*Department of Psychiatry, VA Connecticut Healthcare System, West Haven Campus and Yale University School of Medicine, 911 Campbell Avenue, West Haven, CT 06516*

c*Department of Human Genetics, Emory University School of Medicine, 615 Michael St., Suite 301, Atlanta, GA 30322*

d*Departments of Medicine (Genetics Program), Neurology, Genetics & Genomics, Epidemiology, and Biostatistics, Boston University Schools of Medicine, 715 Albany Street, Boston, MA 02118*

Abstract

The Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA) yields reliable DSM-IV diagnoses for a variety of psychiatric disorders, including alcohol and drug dependence. This study examines the reliability of individual DSM-IV criteria for lifetime substance dependence diagnoses and the impact of those criteria on diagnostic reliability.

Methods—Two hundred ninety-three subjects (52.2% women; 38.2% African American, 46.8% European American, 7.5% Hispanic) were interviewed twice over a two-week period to examine the inter-rater reliability (n = 173) or test-retest reliability (n = 120) of the SSADDA. Cohen's κ statistic and its confidence interval were used to assess the reliability of individual diagnostic criteria.

Results—Overall, the inter-rater reliability estimates were excellent for individual DSM-IV criteria for nicotine and opioid dependence; good for alcohol and cocaine dependence, and fair for dependence on cannabis, sedatives and stimulants. The impact of any individual criterion on diagnostic reliability was minimal, consistent with the notion that the DSM-IV diagnosis of substance dependence measures an underlying construct that is relatively consistent across specific groups of substances.

Conclusions—These results, combined with results from a study of the SSADDA's diagnostic reliability (Pierucci-Lagha et al., 2005), show that the instrument can be used reliably to assess substance dependence.

Keywords

Diagnostic reliability; concordance; DSM-IV criteria; substance dependence; chance-corrected agreement; reliability analysis

*Correspondence: Department of Psychiatry, University of Connecticut Health Center, Farmington, CT 06030-2103, Email: kranzler@psychiatry.uchc.edu, Telephone: 860-679-4151, Fax: 860-679-1316

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.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

1. INTRODUCTION

To enhance the reliability of the psychiatric diagnostic process, particularly as it pertains to psychiatric genetics, several structured or semi-structured interviews have been developed, including the Diagnostic Interview for Genetic Studies (DIGS; Nurnberger et al., 1994), the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA; Bucholz et al., 1995), and the Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA; Pierucci-Lagha et al., 2005).

Bucholz et al. (1995) evaluated the diagnostic reliability of the SSAGA using DSM-III-R criteria (American Psychiatric Association, 1987), in a sample with a high prevalence of alcohol use disorders. The SSAGA also yields DSM-IV diagnoses (American Psychiatric Association, 1994). The SSADDA is a comprehensive psychiatric interview schedule that assesses the physical, psychological, social, and psychiatric manifestations of substance abuse and dependence, and a variety of psychiatric disorders in adults. Similar to, and based on, the SSAGA, the SSADDA aims to increase the validity of the phenotypic information that it provides through the application of a rigorous process of self- and cross-editing. The substantial time required for the interview and for the editing process make the SSADDA useful for the in-depth collection of phenotypic information, particularly suitable for genetic studies, which generally use a cross-sectional phenotypic evaluation. In the course of adapting the SSADDA from the SSAGA, we abridged the alcohol section; expanded the sections on opioid and cocaine dependence; added sections on pathological gambling, attention deficit hyperactivity disorder (ADHD), and environmental factors; and changed the format from paper-and-pencil to computer-assisted administration (see details under the section “Instrument”). The SSADDA provides a more detailed and intensive evaluation of the clinical course of drug dependence and associated psychiatric symptoms and disorders than, for example, the DIGS. Also in contrast to the DIGS, the SSADDA does not require that the interviewer exercise a significant amount of clinical judgment.

Recently, we evaluated the inter-rater and test-retest reliabilites of DSM-IV diagnoses obtained with the SSADDA in a diverse sample of subjects, in which there was a high prevalence of cocaine- and/or opioid-dependent individuals (Pierucci-Lagha et al., 2005). In that study, the SSADDA yielded good-to-excellent reliabilities for most substance dependence diagnoses, although the reliability of substance abuse diagnoses was not as good (Pierucci-Lagha et al., 2005). The present report describes the reliability of the DSM-IV diagnostic criteria obtained using the SSADDA. The reliability of a diagnosis is limited by the reliability of the criteria that comprise it. Information on the reliability of the criteria can assist in revision of the diagnoses, as is regularly done with the Diagnostic and Statistical Manual. Further, as we have demonstrated elsewhere (Gelernter et al., 2005;2006), subtypes of cocaine and opioid dependence derived using diagnostic criteria and other data obtained with the SSADDA can be more useful as traits for linkage analysis than the DSM-IV diagnoses.

In the present study, we evaluated the reliability of individual diagnostic criteria for dependence on the major classes of substances that are evaluated by the SSADDA. In the Bucholz et al. (1995) study, the majority of individual criterion items were reliable, with 87% and 81% showing fair or better reliability for the within- and cross-center studies, respectively. Further, in that study, reliability estimates were good for the alcohol and cocaine items, while those for stimulants were less satisfactory. We also examined the impact of variation in criterion reliability on the reliabilities of substance dependence diagnoses.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

2. METHODS

2.1. Subjects

A total of 293 subjects were each recruited to participate in two separate administrations of the SSADDA. Of this number, 159 subjects were recruited from substance abuse treatment facilities affiliated with the University of Connecticut Health Center (UConn) or the Yale University School of Medicine (Yale). In addition, 59 patients were recruited from inpatient and outpatient psychiatric services at UConn, and 75 community respondents were recruited through advertisements in local media. Inclusion and exclusion criteria are detailed in Pierucci-Lagha et al. (2005). Briefly, subjects were included if they were at least 18 years old and willing to provide informed consent to participate. They were excluded from participation if they showed clinical evidence of a severe psychiatric disorder that might substantially limit their capacity to provide accurate diagnostic information (i.e., schizophrenia, gross cognitive impairment), or if they were unable to read or write English at an 8th grade or higher level. Subjects were paid for their participation. The institutional review boards at UConn and Yale approved the study protocol and informed consent form.

2.2. Instrument

The SSADDA was developed to provide a more detailed coverage of specific drug use disorders, particularly cocaine and opioid use disorders. The SSADDA includes detailed questions on the onset and recency of symptoms for these two major drugs of abuse, and adds some additional items specific to these drugs of abuse (e.g., assessment of symptoms of cocaine-induced paranoia). In addition, sections were added to the SSADDA to cover attention deficit hyperactivity disorder (ADHD) and pathological gambling, both of which are theoretically and clinically relevant to drug dependence (Petry, 2002;Wilens, 2004). Finally, a section on environment covariates considered likely to have an impact on drug and alcohol dependence risk was added.

Administration of the SSADDA is computer assisted, allowing direct entry of subject responses by the interviewer. The computerized format includes core features, such as automatic “skip-outs,” a cross checking function to identify inconsistent responses, a running tabulation of diagnostic criteria, and a check for out-of-range values, features that serve to streamline the interview process and aid in collecting accurate information. The computerized version of the SSADDA also allows for direct upload of data to a database, thereby eliminating the time-consuming step of data entry and verification (with their attendant potential for errors) and permitting the ready generation of DSM-IV diagnoses (American Psychiatric Association, 1994) using scoring algorithms.

As is the case for the SSAGA, the SSADDA allows a trained (non-clinician) interviewer to identify a variety of substance use and psychiatric disorders by collecting information on the onset and recency of symptoms and on their severity and duration. A useful feature of the SSAGA, which was retained in the SSADDA, is its assessment of the relationship between alcohol and drug problem clusters and the occurrence of other psychiatric disorders (Bucholz et al., 1994). Information including the dates of occurrence of alcohol and drug use; clustering of problems; periods of abstinence from alcohol and drug use; and dates of depressive episodes, dysthymic periods, manic episodes, and other psychiatric disorders is correlated so as to categorize the respondent’s history of psychiatric disorders as either completely independent of substance problems, or including at least some psychiatric symptoms that occurred in temporal association with substance use.

Data were scored using algorithms to generate DSM-IV lifetime diagnoses (American Psychiatric Association, 1994). Sources of disagreement, such as variation in the information

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

provided by the subject on the two occasions, were noted. All diagnostic discrepancies, plus a random sample of concordant diagnoses, were rechecked to ensure the accuracy of the data.

2.3. Study design

The study consisted of two sub-studies, focusing on 1) inter-rater reliability and 2) test-retest reliability. Inter-rater reliability was assessed in 173 subjects (i.e., 103 substance abuse patients, 49 psychiatric patients, and 21 community respondents), who were interviewed twice by different interviewers. The purpose of this sub-study was to evaluate the extent to which independent raters can use the same instrument to arrive at the same diagnostic conclusions. To ensure wider applicability of the findings, the inter-rater sub-study was further divided into two components, a within-center component, in which each subject was interviewed by two different interviewers at the same site, and a cross-center component, in which each subject was interviewed once by an interviewer at each of the two participating sites. Test-retest was assessed in 120 subjects (i.e., 56 substance abuse patients, 10 psychiatric patients, and 54 community respondents), who were interviewed twice by the same interviewer, thereby providing a measure of the short-term stability of the diagnosis.

Twelve interviewers participated in this study. All interviewers had a bachelor's degree (three also had a master's degree) and received formal training by the same training coordinator at UConn. All interviews were audiotaped to allow confirmation that the interviewers followed standard procedures, and to provide objective information in the event of coding discrepancies. The coded interviews were scored using algorithms to generate DSM-IV lifetime diagnoses (American Psychiatric Association, 1994) for dependence on the following substances: nicotine, alcohol, cocaine, opioid, cannabis, stimulants, and sedatives.

2.4. Statistical analyses

All analyses included only subjects who reported any lifetime use of the substance, so that the number of subjects varied by substance. Categorical data were analyzed using the κ coefficient, a measure of chance-corrected agreement (Cohen, 1960). Agreement was classified according to a modification of the guidelines developed by Cicchetti and Sparrow (1981), such that a reliability coefficient below 0.40 is considered to be poor; 0.40-0.59 is fair; 0.60-0.74 is good; and 0.75-1.00 is excellent. As reliability estimates are sample dependent, 95% confidence intervals are also provided for κ (Fleiss, 1981).

As in Bucholz et al. (1995), we used two approaches to evaluate the impact of each diagnostic criterion on the reliability of substance dependence diagnoses. First, the reliability of each diagnosis was re-estimated after equating the criterion endorsement in Time 2 to that in Time 1 for each of the seven criteria one at a time. These re-estimated reliability coefficients were then compared with the original estimate in order to examine the impact on diagnostic reliability by changing the endorsement of a single criterion. Second, the number of differences in criterion endorsement among individuals for whom the diagnosis was the same at both times (Stable) were compared with the number of differences among individuals having different diagnoses at the two timepoints (Unstable). These comparisons indicate the number of differences in criterion endorsement that are needed, on average, to change a substance dependence diagnosis. This analysis, unlike the inter-rater and test-retest reliabilities (which used all criteria, irrespective of clustering), was limited to criteria that clustered together during a 12-month period, as required by DSM-IV for all substance dependence diagnoses.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

3. RESULTS

3.1. Subjects and interval between interviews

During the first interview, 2 subjects reported no lifetime use of any of the seven substances examined in the study. The remaining 291 subjects (99.3% of the total) had a mean age of 37.9 years (SD =10.6), and had completed a mean of 13.0 years (SD = 2.4) of education. Nearly half of the subjects (47.1%) were European American, and more than one-third (38.5%) of the sample was African American. The majority of subjects were never married (59.1%). There were no significant differences on any of these demographic variables by sub-study.

The mean and median times elapsed between the two interviews were 13.8 (SD = 5.1) and 13 days, respectively.

3.2. Inter-rater reliability

Table 1 shows the inter-rater reliability estimates (using the κ statistic and its 95% confidence interval) for the individual DSM-IV criteria. Examining the 95% confidence intervals for κ, the vast majority of criteria across the seven groups of substances show satisfactory reliability. The values of κ for nicotine, alcohol, cocaine, and opioid dependence criteria were generally high; none were in the poor range and most were in the good-to-excellent range. Reliability estimates for cannabis dependence symptoms were not as high as those for other substances, although a majority fell in the fair range and only two were in the poor range. Reliability estimates for sedatives and stimulants were less than satifactory for the majority of the items.

3.3. Test-retest reliability

As might be expected, estimates of test-retest reliability were, for the most part, consistent with, but higher than, the estimates of inter-rater reliability (data not shown).

3.4. Impact of criterion reliability on diagnostic reliability

Table 2 shows values of κ recomputed with each criterion, in turn, being treated as reliable in the inter-rater sub-study. κ estimates were very similar to the original estimates for all substances, and in no case was the difference significant, as reflected by the overlapping confidence intervals. Similar findings were obtained for the test-retest sub-study (data not shown).

As is shown in Table 3 for the inter-rater sub-study, across all substances, diagnoses were generally stable. In this sub-study, opioid dependence showed the lowest rate of instability (6.8%) and cannabis dependence the highest (19.0%). In the test-retest sub-study (data not shown), similar findings obtained, except that nicotine dependence showed the lowest rate of instability (2.0%) and sedative dependence showed the highest rate (13.6%). For unstable diagnoses, in both sub-studies, the mean number of criteria changes exceeded 3 for most substances.

4. DISCUSSION

This study examined the inter-rater and test-retest reliability of the DSM-IV criteria for dependence on nicotine, alcohol, cocaine, opioid, cannabis, stimulants and sedatives. Using Cohen's κ, we found fair-to-excellent reliability for the criteria for nicotine, alcohol, cocaine, opioid, and cannabis dependence, while stimulant and sedative dependence criteria were less reliable. Of the 49 criteria evaluated in the inter-rater sub-study, 12 (24.5%) showed poor reliability (i.e., $\kappa < 0.40$). In the test-retest sub-study, 8 criteria (16.3%) showed poor reliability. In both sub-studies, the vast majority of unreliable criteria were for stimulants and sedative dependence, however, the smaller number of users of these substances, resulting in lower

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

stability of κ, may have contributed to the lower estimates of reliability for these criteria. These findings are similar to those obtained in the within- and cross-center reliability studies of the SSAGA, in which there were 13% and 19%, respectively, of criteria that were found to be unreliable (Bucholz et al., 1995).

With respect to individual substances, all but one opioid dependence criterion (i.e., "using larger amounts or over a longer period than intended") yielded a value of κ that was good or better (i.e., $\kappa \geq 0.60$). Opioid dependence criteria evaluated using the SSAGA also showed high reliability (Bucholz et al., 1995). Similarly high estimates of reliability were obtained for nicotine dependence criteria using both the SSADDA and the SSAGA. Values of κ for tolerance were either fair or lower, except for opioids and sedatives. In contrast, values of κ for withdrawal were good to excellent except for cannabis, stimulants, and sedatives. Although tolerance and withdrawal are both manifestations of neuroadaptation, it is likely that they have different underlying mechanisms (Cicero, 1980;Langenbucher et al., 1997). In addition, withdrawal is based on more dramatic and perhaps more reliably reported symptoms than is tolerance, a characteristic that could be important when establishing a history of the phenomena retrospectively. Finally, the problem with tolerance might relate to the definition used; for example, a 50% increase in the amount of the substance needed for an effect might be a relatively low threshold (Schuckit et al., 1999).

The reliability estimates for individual SSADDA alcohol criteria were generally good, though somewhat lower than those obtained using the SSAGA (Bucholz et al., 1995). Although reliability estimates were good for most cocaine dependence criteria when elicited using the SSADDA, values of κ for these criteria were also somewhat lower than those reported using the SSAGA (Bucholz et al., 1995).

In the present study, reliability for most cannabis dependence criteria were fair, similar to that observed in other studies (Compton et al., 1996;Cottler et al., 1997;Miele et al., 2000;Bucholz et al., 1995). These findings complement the growing body of literature supporting the existence of the cannabis dependence syndrome and cannabis withdrawal, as well as the validity of the DSM IV dependence criteria as they apply to cannabis dependence (Budney, 2006;Budney and Hughes, 2006). The findings of fair reliability for cannabis withdrawal criteria, despite the heterogeneity of the sample, complement recent findings showing that cannabis withdrawal is both clinically significant and readily identified even in non-treatment seeking populations (Copersino et al., 2006). Reliability estimates for stimulants and sedatives were less than satisfactory, similar to those reported for the SSAGA (Bucholz et al., 1995), perhaps due in both studies to the relatively low frequency of endorsement for criteria for these disorders.

Substance-dependent subjects in the present study sample predominanly included individuals with opioid and/or cocaine dependence (cf. Gelernter et al., 2005;2006), while the SSAGA substance-dependent study sample predominantly included individuals with alcohol dependence. We also evaluated reliability in a more heterogeneous sample than that studied by Bucholz et al. (1995). Specifically, subjects in our sample had a mean of 3.1 substance use disorder diagnoses, with 34.4% having four or more such diagnoses (Pierucci-Lagha et al., 2005). Although comparable numbers are not reported by Bucholz et al. (1994,1995), it is evident from the data presented that their reliability study sample was predominantly comprised of individuals with alcohol use disorders and with fewer comorbid substance use disorders than in the subjects we studied. In the context of multiple substance use disorders, subjects are more likely to misattribute their symptoms to one substance rather than another, leading to reduced criterion reliability for some disorders.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

It should be noted also that estimates of κ are not directly comparable across the SSADDA and SSAGA studies, as the populations evaluated and the prevalence of the disorders differed substantially between them. Although the κ statistic shows the extent to which agreement exceeds that predicted by chance alone, it is not possible directly to interpret differences in κ across different conditions in different studies (Stemler, 2004). Administration of the SSAGA and SSADDA to the same subjects under controlled conditions would be required for such a direct comparison. Interpretation of reliability findings should, then, be done in the context of a particular study sample.

The results of the impact study conducted using the SSADDA suggest that the diagnosis of dependence on different substances did not generally depend on a single criterion. Specifically, none of the seven DSM-IV criteria for substance dependence individually influenced diagnostic reliability; rather, it was necessary for at least three criteria on average to change in order to affect the reliability of a substance dependence diagnosis. These findings are similar to those of Bucholz et al. (1995).

In summary, these findings showing the reliability of the individual DSM-IV criteria for dependence on different substances provides additional support for the reliability of the SSADDA in the evaluation of substance dependence. Based on its reliability in the assessment of individual criteria and diagnoses (Pierucci-Lagha et al., 2005), its poly-diagnostic nature, its suitability for administration by lay interviewers, and its computer-assisted format, which includes skip patterns and internal consistency checking, the SSADDA is a useful diagnostic instrument for a variety of applications, including genetic and family studies of substance dependence.

ACKNOWLEDGEMENTS

Supported by NIH grants DA12422, DA12849, DA12890, DA15105, DA15766, and AA13736. The authors thank the COGA investigators for generously allowing the SSAGA interview to be modified for use by our research group, Jennifer Hamilton and John Farrell for developing the electronic version of the SSADDA, and Deborah Cebrik for database design and management. The authors also thank all of the interviewers for their contributions to the study.

REFERENCES

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 3rd ed., revised. American Psychiatric Press; Washington DC: 1987.

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th ed.. American Psychiatric Press; Washington DC: 1994.

Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JI Jr. 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. J Stud Alcohol 1994;55:149–158. [PubMed: 8189735]

Bucholz KK, Hesselbrock VM, Shayka JJ, Nurnberger JI Jr. Schuckit MA, Schmidt I, Reich T. Reliability of individual diagnostic criterion items for psychoactive substance dependence and the impact on diagnosis. J Stud Alcohol 1995;56:500–505. [PubMed: 7475029]

Budney AJ. Are specific dependence criteria necessary for different substances: how can research on cannabis inform this issue? Addiction 2006;101(Suppl 1):125–133. [PubMed: 16930169]

Budney AJ, Hughes JR. The cannabis withdrawal syndrome. Curr Opin Psychiatry 2006;19:233–238. [PubMed: 16612207]

Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. Am J Ment Defic 1981;86:127–137. [PubMed: 7315877]

Cicero, TJ. Alcohol self-administration, tolerance, and withdrawal in humans and animals: theoretical and methodological issues. In: Rigter, H.; Rabbe, JCC., editors. Alcohol Tolerance and Dependence. Elsevier; New York: 1980. p. 1-51.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20:37–46.

ComptonWMCottlerLBDorseyKBSpitznagelELMagerDEComparing assessments of DSM-IV substance dependence disorders using CIDI-SAM and SCAN. Drug Alcohol Depend199641179187 Erratum in: Drug Alcohol Depend. 1996, 42, 217-219 [PubMed: 8842630]

Copersino ML, Boyd SJ, Tashkin DP, Huestis MA, Heishman SJ, Dermand JC, Simmons MS, Gorelick DA. Cannabis withdrawal among non-treatment-seeking adult cannabis users. Am J Addict 2006;15:8–14. [PubMed: 16449088]

Cottler LB, Grant BF, Blaine J, Mavreas V, Pull C, Hasin D, Compton WM, Rubio-Stipec M, Mager D. Concordance of DSM-IV alcohol and drug use disorder criteria and diagnoses as measured by AUDADIS-ADR, CIDI and SCAN. Drug Alcohol Depend 1997;47:195–205. [PubMed: 9306045]

Fleiss, JL. Statistical methods for rates and proportions. 2nd edition. John Wiley & Sons, Inc; New York: 1981.

Gelernter J, Panhuysen C, Weiss R, Brady K, Hesselbrock V, Rounsaville B, Poling J, Wilcox M, Farrer L, Kranzler HR. Genomewide linkage scan for cocaine dependence and related traits: Linkages for a cocaine-related trait and cocaine-induced paranoia. Am J Med Genet Part B (Neuropsychiatric Genetics) 2005;136B:45–52.

Gelernter J, Panhuysen C, Wilcox M, Hesselbrock V, Rounsaville B, Poling J, Weiss R, Sonne S, Farrer L, Kranzler HR. Genomewide linkage scan for opioid dependence and related traits. Am J Hum Genet 2006;78:759–769. [PubMed: 16642432]

Langenbucher J, Chung T, Morgenstern J, Labouvie E, Nathan PE, Bavly L. Physiological alcohol dependence as a “specifier” of risk for medical problems and relapse liability in DSM-IV. J Stud Alcohol 1997;58:341–350. [PubMed: 9203114]

Miele GM, Carpenter KM, Smith Cockerham M, Trautman KD, Blaine J, Hasin DS. Substance Dependence Severity Scale (SDSS): reliability and validity of a clinician-administered interview for DSM-IV substance use disorders. Drug Alcohol Depend 2000;59:63–75. [PubMed: 10706976]

Nurnberger JI Jr. Blehar MC, Kaufmann CA, York-Cooler C, Simpson SG, Harkavy-Friedman J, Severe JB, Malaspina D, Reich T. Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. Arch Gen Psychiatry 1994;51:849–859. [PubMed: 7944874]

Petry NM. How treatments for pathological gambling can be informed by treatments for substance use disorders. Exp Clin Psychopharmacol 2002;10:184–192. [PubMed: 12233980]

Pierucci-Lagha A, Gelernter J, Feinn R, Cubells JF, Pearson D, Pollastri A, Farrer L, Kranzler HR. Diagnostic reliability of the Semi-structured Assessment for Drug Dependence and Alcoholism (SSADDA). Drug Alcohol Depend 2005;80:303–312. [PubMed: 15896927]

Schuckit MA, Daeppen JB, Danko GP, Tripp ML, Smith TL, Li TK, Hesselbrock VM, Bucholz KK. Clinical implications for four drugs of the DSM-IV distinction between substance dependence with and without a physiological component. Am J Psychiatry 1999;156(1):41–49. [PubMed: 9892296]

StemlerSEA comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. Practical Assessment, Research & Evaluation200494 Retrieved December 14, 2006 from http://PAREonline.net/getvn.asp?v=9&n=4

Wilens TE. Attention-deficit/hyperactivity disorder and the substance use disorders: the nature of the relationship, subtypes at risk, and treatment issues. Psychiatr Clin North Am 2004;27:283–301. [PubMed: 15063998]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 1

Inter-rater reliability of individual criteria for specific DSM-IV substance dependence diagnoses

DSM-IV Criteria	Nicotine (n = 152)	Alcohol (n = 171)	Cocaine (n = 110)	Opioid (n = 59)	Cannabis (n = 142)	Stimulants (n = 50)	Sedatives (n = 27)
(1) Tolerance	0.55 (0.41, 0.68)	0.54 (0.42, 0.66)	0.49 (0.33, 0.65)	0.62 (0.42, 0.82)	0.55 (0.38, 0.71)	0.34 (-0.02, 0.70)	0.60 (0.19, 1.00)
(2) Withdrawal	0.65 (0.52, 0.77)	0.66 (0.53, 0.79)	0.60 (0.44, 0.75)	0.90 (0.79, 1.00)	0.45 (0.27, 0.64)	0.39 (-0.04, 0.82)	0.47 (-0.13, 1.00)
(3) Taking the substance in larger amounts or over a longer period than was intended	0.64 (0.52, 0.77)	0.65 (0.53, 0.76)	0.46 (0.27, 0.64)	0.56 (0.34, 0.78)	0.53 (0.38, 0.68)	0.23 (-0.22, 0.69)	0.13 (-0.30, 0.56)
(4) Persistent desire, or unsuccessful efforts to cut down or control substance use	0.75 (0.65, 0.86)	0.63 (0.51, 0.74)	0.47 (0.25, 0.69)	0.69 (0.50, 0.88)	0.43 (0.28, 0.59)	0.12 (-0.24, 0.47)	-0.07 (-0.15, 0.05)
(5) Great deal of time is spent in activities necessary to obtain, use, or recover from substance use	0.76 (0.66, 0.87)	0.53 (0.39, 0.68)	0.56 (0.39, 0.73)	0.73 (0.55, 0.90)	0.36 (0.19, 0.53)	0.30 (-0.20, 0.79)	-0.06 (-0.17, 0.04)
(6) Giving up or reducing important social, occupational, or recreational activities	0.53 (0.29, 0.76)	0.70 (0.58, 0.81)	0.58 (0.42, 0.73)	0.71 (0.52, 0.90)	0.44 (0.26, 0.62)	-0.06 (-0.13, 0.01)	0.51 (0.03, 0.99)
(7) Continuing substance use despite persistent, or recurrent physical or psychological problems	0.76 (0.65, 0.87)	0.58 (0.45, 0.72)	0.53 (0.35, 0.71)	0.79 (0.63, 0.95)	0.30 (0.09, 0.51)	0.50 (0.16, 0.85)	0.26 (-0.18, 0.71)

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 2

Impact of individual criteria on inter-rater reliability of specific DSM-IV substance dependence diagnoses

Original or DSM-IV Dependence Criteria	Estimated κ and 95% Confidence Interval						
	Nicotine (n = 152)	Alcohol (n = 171)	Cocaine (n = 110)	Opioid (n = 59)	Cannabis (n = 142)	Stimulants (n = 50)	Sedatives (n = 27)
Original	0.75 (0.65, 0.86)	0.66 (0.54, 0.77)	0.59 (0.40, 0.77)	0.86 (0.73, 0.99)	0.52 (0.36, 0.68)	0.43 (0.05, 0.81)	0.51 (0.03, 0.99)
(1)	0.76 (0.66, 0.87)	0.67 (0.55, 0.78)	0.62 (0.44, 0.80)	0.86 (0.73, 0.99)	0.52 (0.37, 0.68)	0.43 (0.05, 0.81)	0.51 (0.03, 0.99)
(2)	0.74 (0.63, 0.84)	0.68 (0.57, 0.79)	0.58 (0.40, 0.77)	0.90 (0.78, 1.00)	0.55 (0.39, 0.70)	0.43 (0.05, 0.81)	0.51 (0.03, 0.99)
(3)	0.80 (0.71, 0.90)	0.67 (0.55, 0.78)	0.60 (0.42, 0.78)	0.90 (0.78, 1.00)	0.53 (0.38, 0.69)	0.43 (0.05, 0.81)	0.51 (0.03, 0.99)
(4)	0.76 (0.66, 0.87)	0.63 (0.51, 0.75)	0.57 (0.38, 0.76)	0.86 (0.73, 0.99)	0.54 (0.38, 0.70)	0.43 (0.05, 0.81)	0.51 (0.03, 0.99)
(5)	0.72 (0.62, 0.83)	0.66 (0.54, 0.77)	0.57 (0.38, 0.76)	0.86 (0.73, 0.99)	0.50 (0.34, 0.66)	0.43 (0.05, 0.81)	0.51 (0.03, 0.99)
(6)	0.76 (0.66, 0.87)	0.68 (0.57, 0.79)	0.53 (0.34, 0.72)	0.86 (0.73, 0.99)	0.51 (0.35, 0.67)	0.43 (0.05, 0.81)	0.51 (0.03, 0.99)
(7)	0.74 (0.63, 0.84)	0.67 (0.56, 0.78)	0.57 (0.38, 0.76)	0.90 (0.78, 1.00)	0.54 (0.38, 0.70)	0.43 (0.05, 0.81)	0.51 (0.03, 0.99)

Drug Alcohol Depend. Author manuscript; available in PMC 2008 November 2.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 3

Number of diagnostic criteria items changes by stability of substance dependence diagnoses for inter-rater reliability substudy

Dependence Diagnosis Stability	Sample Size Mean ± Standard Deviation						
	Nicotine	Alcohol	Cocaine	Opioid	Cannabis	Stimulants	Sedatives
Stable	(n = 133) 0.84 ± 1.16	(n = 143) 0.63 ± 1.17	(n = 95) 1.35 ± 1.40	(n = 55) 0.55 ± 1.03	(n = 115) 0.58 ± 1.26	(n = 44) 0.16 ± 0.75	(n = 24) 0.21 ± 0.72
Unstable	(n = 19) 3.68 ± 1.06	(n = 28) 3.86 ± 1.21	(n = 15) 4.20 ± 1.15	(n = 4) 3.25 ± 0.50	(n = 27) 4.15 ± 1.20	(n = 6) 4.83 ± 1.17	(n = 3) 5.00 ± 2.00