Reliability and Validity of the Alcohol Short Index of Problems and a Newly Constructed Drug Short Index of Problems*

ARTHUR I. ALTERMAN, PH.D.,† JOHN S. CACCIOLA, PH.D.,† MEGAN A. IVEY, M.S., M.L.A., BRIAN HABING, PH.D.,† AND KEVIN G. LYNCH, PH.D.

Center for Studies of Addiction, University of Pennsylvania, 3440 Market Street, Suite 370, Philadelphia, Pennsylvania 19104

ABSTRACT. Objective: This study evaluated the psychometric properties of the 15-item alcohol Short Index of Problems (SIP) instrument and those of a newly constructed 15-item drug Short Index of Problems (SIP-D) instrument in 277 newly entered substance-abuse patients. **Method:** The SIP is derived from the longer, 50-item Drinker Inventory of Consequences (DrInC), which was designed to assess adverse consequences of alcohol use. The SIP-D was constructed by substituting the term "drug use" for the term "drinking" in each SIP item. A 3-month recall interval was employed. **Results:** Factor analyses of each of the instruments revealed similar solutions, with only one main factor accounting for the majority of variance. Nonparametric item response theory methods produced the same finding. Internal consistency reliability estimates for the SIP and SIP-D total scores were .98 and .97, respectively. Concurrent validity was demonstrated by examining the correlations of the total scores for each of the instruments with the recent summary indexes of the newly revised Addiction Severity Index (ASI-Version 6): alcohol, drug, medical, economic, legal, family/social, and psychiatric problems. **Conclusions:** This study is the first to confirm the psychometric validity of the SIP when used as an independent instrument unembedded within the DrInC. The study also supports the use of the SIP-D as a brief measure of adverse consequences of drug use. The findings strongly support the unidimensional structure of both measures. (*J. Stud. Alcohol Drugs* **70:** 304-307, 2009)

EFFORTS TO ASSESS ADVERSE CONSEQUENCES of alcohol and drug use often have been confounded with symptoms of dependence and quantity and frequency of substance use (Blanchard et al., 2003). Indeed, research suggests that consumption levels alone are not necessarily good predictors of substance use-related impairment (Bender et al., 2007). Within this context, Miller and colleagues (1995) developed the Drinking Inventory of Consequences (DrInC) for alcohol-dependent patients to serve as a relatively brief, self-report measure of the severity of the consequences of alcohol use. Constructed by a panel of experts, the DrInC consists of 45 primary items in five domains: Physical, Intrapersonal, Social Responsibility, Interpersonal, and Impulse Control consequences. The instrument can be administered using either a lifetime or a 3-month recall timeframe. The DrInC also includes an additional five "control" items designed to detect careless responding or dishonesty. These items are not included in psychometric evaluations of the instrument. Although factor analysis has not supported five independent domains (Anderson et al., 2001; Miller et al. 1995), these domains have been used in research in addition to the total score. The generally excellent internal consistency reliability and test-retest reliability of each of the subscales appear to have been accepted as indications of their independent integrity.

Miller et al. (1995) also developed a short version of the DrInC labeled the Short Index of Problems (SIP). The SIP was derived by including the three items in each DrInC subscale that were most highly correlated with the total subscale score. Internal consistency of this new instrument was lower than for the full DrInC subscales, ranging from .57 to .66; that for the entire instrument was .81. Six-month test reliability was good for both the subscales and the total score. In another study that used the full DrInC (Feinn et al., 2003), the internal consistency of the embedded 3-month SIP subscales ranged from .56 to .64, with test-retest reliability ranging from .17 to .47 when the more conservative intraclass correlation (ICC) statistic was employed. Confirmatory factor analyses (CFAs) of the SIP failed to yield a model with adequate fit.

To construct a battery of external, concurrent validity measures for a project revising the Addiction Severity Index (ASI-6; McLellan et al., 2006), we included the 15-item, 3-month SIP as a freestanding external validity measure of recent alcohol problems. We also constructed a drug SIP instrument (SIP-D) by changing the reference to "drinking" in the original SIP's (henceforth designated SIP-A) instructions and individual items to "drug use." This approach was

Received: May 12, 2008. Revision: August 20, 2008.

*This research was supported by National Institute on Drug Abuse center grants P50 awarded to A. Thomas McLellan and P60 awarded to Charles P. O'Brien.

†Correspondence may be sent to Arthur I. Alterman at the above address or via email at: alterman@mail.trc.upenn.edu. John S. Cacciola is with the Treatment Research Institute, Philadelphia, PA. Brian Habing is with the University of South Carolina, Columbia, SC.

similar to that used in the construction of the Inventory of Drug Use Consequences (InDUC), a variant of the DrInC described in the original Miller et al. (1995) work, which added "or drugs" after "drinking" in the DrInC items to obtain a measure of the adverse consequences of alcohol *or* drugs. Although Miller et al. (1995) did not report psychometric data for the InDUC, a more recent study by Tonigan and Miller (2002) generally found good 2-day test-retest reliability (ICC) for four of the five subscales of the lifetime InDUC in a small sample of heroin or cocaine misuse patients (N = 27-30). In the same article, CFAs were conducted on the data of a larger sample (N = 191) of polysubstance abusers to investigate the fits of four- and five-factor models. The four-factor model fit significantly better than the five-factor model. However, tests of fit for models with fewer factors were not evaluated. Blanchard et al. (2003) administered the 3-month version of the InDUC to substance misusers. An exploratory factor analysis (EFA) of the InDUC's items yielded one primary factor. Using the 15 items in the InDUC with the strongest item-total correlations, rather than those with the strongest item-subscale correlations, these investigators also created an abbreviated version of the InDUC, entitled the SIP-AD. The correlation between the SIP-AD and the InDUC was .96, and the internal consistency of the SIP-AD total score was .95.

In another study, Bender et al. (2007) administered a modified version of the 3-month SIP-A (SIP-SUD) to outpatient psychoactive substance use-disordered patients with concurrent bipolar disorder. The term "or drug use" was appended after "drinking" in each item, as had been done for the InDUC. Cronbach α's for the five subscales ranged from .62 to .88 and that for the total SIP-SUD was .93. One-week test-retest reliability correlations also were good. An EFA of the SIP-SUD yielded two factors, a main factor explaining 49.9% of the variance comprising 14 of the 15 items and a second factor (10.2% of the variance) comprising one item, "I had an accident while drinking or using drugs." Thus, the research done with the InDUC or its abbreviated versions suggests that altering the wording of the original DrInC items to assess the effects of drug use as well as alcohol use resulted in generally reliable and valid products. Although the InDUC and its abbreviated versions have the advantage of assessing the effects of either alcohol or drug use, there often is a need to assess just the specific effects of drug use, as embodied in the SIP-D.

The objectives of the present research were to evaluate the psychometric properties of both the SIP-A and the SIP-D. Relatively little research has described the reliability and validity of the SIP-A, and some of the work studied the SIP-A embedded in the full DrInC instrument (Feinn et al., 2003; Miller et al., 1995)—which is problematic, because context can modify the effects of an instrument (Alterman et al., 2003). Thus, in this report we examine the psychometric properties of the 15-item SIP-A. A second research objective was to evaluate the psychometric properties of the newly developed drug-specific short form, the SIP-D.

Method

Participants

The study participants were 277 psychoactive substance use-disordered outpatients who were administered a concurrent validity battery along with the revised ASI, the ASI-6. These participants had recently entered into one of two outpatient treatment programs, two methadone maintenance programs, or two inpatient treatment programs. Three of the programs were publicly funded community programs, and three were Veterans Administration programs. All participants provided informed consent. The research was approved by institutional review boards of the University of Pennsylvania, the City of Philadelphia, and the Veterans Administration.

Eighty-one percent of the study sample were men, the mean (SD) subject age was 41.3 (11.1) years, and 55.4% were black. The average subject had completed more than 11 years of schooling (mean = 11.3 [2.1]), 11% were currently married or living as married, 11.5% were currently working full or part time, and 46% had resided in a homeless shelter in their lifetime. One in four (25.5%) had been legally mandated to treatment. The participants reported a mean of 5.8 (5.6) substance-misuse treatment episodes. The primary substances of use reported by the sample were heroin (36.3%), cocaine/crack (32.4%), alcohol (18.3%), and marijuana (5.0%).

Assessments

All participants were administered the revised ASI-6 semistructured interview within the first 2 weeks of treatment, as well as the SIP-A and SIP-D. The ASI-6 revision of the ASI-5 retains many of the features of the earlier instrument, including assessment of sociodemographic information and the multidimensional assessment of seven problem areas, among which were drug, alcohol-related, medical, employment/economic, legal, family/social, and psychiatric problems. The ASI-6 uses a 30-day timeframe in querying for recent problems that are reported in this study. The revision expanded on the ASI-5 by obtaining more information on current living situation, problems with children, and medical- and economic-related problems. It also updated the instrument by including items on gambling, smoking, and AIDS (McLellan et al., 2006).

Analyses

CFA was conducted on each of the instruments to ascertain whether the purported five-factor structure was present.

Additionally, EFA using the Varimax rotation was conducted on all instruments to further examine their underlying factor structures, and alphas were derived for each derived factor as an indication of the factor's internal consistency. Given that each item has only four response categories, with more than 50% of responses being in an extreme category, the normality assumptions for the standard CFA based on the Pearson product moment correlation seemed untenable. Instead, the analyses were conducted using the polychoric correlations as implemented in MPlus (Muthén and Muthén, 1998-2007). The default WLSMV (weighted least squares mean and variance adjusted) and the ULS (unweighted least squares) estimation methods were used for the CFA and the EFA, respectively. Following previous work on the ASI-5 (Alterman et al., 2007), the nonparametric item response theory-based Mokken scaling (e.g., Sijtsma and Molenaar, 2002) and conditional covariance methods (Stout et al., 1996) were used to provide additional confirmation that each instrument essentially measured a single construct. Concurrent validity was ascertained by examining the correlations between each version of the SIP and the recent problem summary scores for the revised ASI-6's problem areas.

Results

Factor analytic findings

The CFAs of the SIP-A and the SIP-D resulted in estimated correlation matrices for proposed underlying factors having negative eigenvalues. The estimated interfactor correlations ranged from .91 to more than 1.00 for the SIP-A and from .92 to .98 for the SIP-D. In both cases, the interfactor correlations were very similar, whether all five factors were analyzed at once or whether they were analyzed in a pairwise fashion to avoid the negative eigenvalues. This result indicates that the purported five-factor structure failed to fit the data, owing to the five factors not representing distinct underlying constructs. The EFA indicated a one-factor solution as being optimal for the SIP-A. The first eigenvalue was 13.29 (88.6% of the correlation structure explained), with the second being only .53. As shown in Table 1, all 15 items loaded significantly on this factor, with 14 of the 15 items with loadings exceeding .90. The item "Have had an accident while drinking or intoxicated" had a lower loading (.70) on this factor (see Bender et al., 2007). The α coefficient for the SIP-A factor was .98.

A very similar one-factor solution was obtained for the SIP-D. The largest eigenvalue was 12.37, accounting for 82.4% of the correlation structure. Thirteen of the 15 items loaded on this factor with values > .84. The "spent too much time because of" item had a loading of .66 and the "had an accident while" item had a lower but still sizable loading of .43. The α coefficient for this factor was .97.

Table 1. Exploratory factor analysis findings for the SIP-A and SIP-D items

Item	SIP-A Loading	SIP-D Loading
I have been unhappy because of ...	.95	.94
I have not eaten properly because of …	.93	.90
Failed to do what was expected because of …	.96	.94
Felt guilty because of …	.95	.95
Taken foolish risks because of …	.95	.96
Done impulsive things when …	.94	.91
Physical health harmed by …	.92	.84
Had money problems because of …	.96	.94
Physical appearance harmed by …	.93	.88
Family hurt by …	.97	.96
Friendship damaged by ….	.94	.90
… gotten in the way of my growth	.98	.94
… damaged my social life	.96	.92
Spent too much time because of …	.96	.66
Have had an accident while …	.70	.43
Eigenvalue	13.29	12.37
% of variance	88.6	82.4
α	.98	.97

Notes: SIP-A = Short Index of Problems-Alcohol; SIP-D = Short Index of Problems-Drugs.

Nonparametric item response-theory findings

The Mokken scaling analysis of the SIP-A revealed that the items formed a strong (i.e., values > .495) scale (H = .84, minimum H_i = .66) that would be appropriately summarized by a single score. A similar result was obtained for the SIP-D (H = .79, minimum H_i = .55). When used to forcibly split the SIP-A and the SIP-D into optimal potential subscales, the conditional covariance methods chose two parallel subscale solutions for both the SIP-A and the SIP-D (Items 1-6, 15 as one subscale; 7-9, 11-14 as the other, with Item 10 having uncertain assignment). These candidate subscales had attenuation-adjusted estimated correlations of .96 for the SIP-A and .95 for the SIP-D, indicating that any multidimensionality present in the scale is extremely weak. This finding again supported treating each instrument as being summarized by a single score.

Concurrent validity

The correlation between the SIP-A and the ASI-6's alcohol problem summary score was found to be .68. The corresponding correlation between the SIP-D and the ASI-6's drug problem summary score was .61. The correlations between the SIP-A score and the ASI-6's drug, medical, employment/economic, legal, family/social, and psychiatric problems recent summary scores were -.14, .15, -.06, -.04, .17, and .26, respectively. Similarly, the correlations between the SIP-D score and the ASI-6 alcohol, medical, employment/economic, legal, family/social, and psychiatric problems recent summary scores were .01, .19, -.03, .28, .25, and .37. The relatively elevated correlations between the two

SIP scores and ASI psychiatric problems are consistent with the multidimensional problems of psychoactive substance use-disordered patients. The higher correlation of the SIP-D to legal problems than for the SIP-A is understandable, given the illegality/legality of drug use versus alcohol use in our society. As an additional validity evaluation, the SIP-A score was correlated with an individual ASI-6 item that described the period of peak use of alcohol during the past 6 months. This correlation was .66. The SIP-D score was correlated in parallel fashion with an ASI-6 item that described the period of peak use of illicit drug use during the past 6 months. This correlation was .67. As another form of comparison, the correlation between the SIP-A and the SIP-D was evaluated and determined to be .19. The correlation between the ASI-6's drug and alcohol summary measures was -.05.

Discussion

The findings of this study provided support for the reliability and concurrent validity of the alcohol SIP (SIP-A) as well as of the drug SIP (SIP-D), which was developed specifically for this study. The SIP-A was found to yield one primary factor—a finding consistent with prior work with the 45-item parent DrInC (Anderson et al., 2001), the SIP (Feinn et al., 2003), and two abbreviated versions of the InDUC (Bender et al., 2007; Blanchard et al., 2003). Similar findings were obtained for the SIP-D. These findings, taken as a body, indicate that there is no statistical justification for employing separate subscales for these instruments.

Because the SIP-A was administered as a separate instrument in its 15-item format, the findings take on additional value, as we are unaware of any prior study that has focused on the psychometric properties of the SIP that used it unembedded within the DrInC. The α coefficient of the SIP-A was extremely high (.98), and its validity was confirmed by a strong correlation with the ASI-6's alcohol summary score as well as other analyses. Similar findings were obtained for the SIP-D.

As noted, there may be advantages in some circumstances in having an instrument such as the InDUC or its abbreviated versions that can assess the adverse consequences of either alcohol or substance use. At the same time, there also would appear to be benefits in obtaining information about the separate consequences of alcohol use and drug use. In demonstrating high reliability and concurrent validity for the SIP-D, as well as evidence for its status as a unitary construct, the findings obtained are consistent with those obtained by the bulk of the previous work on the DrInC/InDUC and their abbreviated variants. These findings provide convincing support that the SIP-D has value as a brief, self-report assessment to specifically determine the adverse consequences of substance use.

At the same time, we encourage further psychometric research on both the SIP-A and the SIP-D. One limitation of the current study was that test-retest data were not available. Furthermore, although the sample size was adequate for conducting the primary analyses, it was insufficient for the determination of generalizability to sociodemographic subgroups.

References

Alterman, A.I., Cacciola, J.S., Habing, B., and Lynch, K.G. Addiction Severity Index recent and lifetime summary indexes based on nonparametric item response theory methods. Psychol. Assess. **19:** 119-132, 2007.

Alterman, A.I., McDermott, P.A., Cacciola, J.S., Mulvaney, F.D., Rutherford, M.J., Searles, J.S., Lynch, K., and Cnaan, A. Comparison of embedded and isolated administrations of the California Personality Inventory's Socialization Subscale. Psychol. Assess. **15:** 64-70, 2003.

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), Washington, DC, 1994.

Anderson, B.J., Gogineni, A., Charuvastra, A., Longabaugh, R., and Stein, M.D. Adverse drinking consequences among alcohol abusing intravenous drug users. Alcsm Clin. Exp. Res. **25:** 41-45, 2001.

Bender, R.E., Griffin, M.L., Gallop, R.J., and Weiss, R.D. Assessing negative consequences in patients with substance use and bipolar disorders: Psychometric properties of the Short Inventory of Problems (SIP). Amer. J. Addict. **16:** 503-509, 2007.

Blanchard, K.A., Morgenstern, J., Morgan, T.J., Labouvie, E.W., and Bux, D.A. Assessing consequences of substance use: Psychometric properties of the Inventory of Drug Use Consequences. Psychol. Addict. Behav. **17:** 328-331, 2003.

Feinn, R., Tennen, H., and Kranzler, H.R. Psychometric properties of the Short Index of Problems as a measure of recent alcohol-related problems. Alcsm Clin. Exp. Res. **27:** 1436-1441, 2003.

McLellan, A.T., Cacciola, J.C., Alterman, A.I., Rikoon, S.H., and Carise, D. The Addiction Severity Index at 25: Origins, contributions, and transitions. Amer. J. Addict. **15:** 113-124, 2006.

Miller,W.R., Tonigan, J.S., and Longabaugh, R. The Drinker Inventory of Consequences (DrInC): An Instrument for Assessing Adverse Consequences of Alcohol Abuse (Test Manual). NIAAA Project MATCH Monograph Series, Vol. 4, NIH Publication No. 95-3911, Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism, 1995.

Muthén, L.K. and Muthén, B.O. Mplus User's Guide, Version 5, Los Angeles, CA: Muthén and Muthén, 1998-2007.

Sijtsma, K. and Molenaar, I.W. Introduction to Nonparametric Item Response Theory, Thousand Oaks, CA: Sage, 2002.

Stout, W., Habing, B., Douglas, J., Kim, H.R., Roussos, L., and Zhang, J. Conditional covariance-based nonparametric multidimensionality assessment. Appl. Psychol. Meas. **20:** 331-354, 1996.

Tonigan, J.S. and Miller, W.R. The Inventory of Drug Use Consequences (InDUC): Test-retest stability and sensitivity to detect change. Psychol. Addict. Behav. **16:** 165-168, 2002.