



Published in final edited form as:

Drug Alcohol Depend. 2009 January 1; 99(1-3): 18–27. doi:10.1016/j.drugalcdep.2008.06.012.

Validation of the Spanish Addiction Severity Index Multimedia Version (S-ASI-MV)

Stephen F. Butler^{a,†}, José Pedro Redondo^b, Kathrine C. Fernandez^a, and Albert Villapiano^a

^a*Inflexxion, Inc., Newton, MA 02464*

^b*Institute for Family Development, Miami, FL 33144*

Abstract

This study aimed to develop and test the reliability and validity of a Spanish adaptation of the ASI-MV, a computer administered version of the Addiction Severity Index, called the S-ASI-MV. Participants were 185 native Spanish-speaking adult clients from substance abuse treatment facilities serving Spanish-speaking clients in Florida, New Mexico, California, and Puerto Rico. Participants were administered the S-ASI-MV as well as Spanish versions of the general health subscale of the SF-36, the work and family unit subscales of the Social Adjustment Scale Self-Report, the Michigan Alcohol Screening Test, the alcohol and drug subscales of the Personality Assessment Inventory, and the Hopkins Symptom Checklist-90. Three-to-five-day test-retest reliability was examined along with criterion validity, convergent/discriminant validity, and factorial validity. Measurement invariance between the English and Spanish versions of the ASI-MV was also examined. The S-ASI-MV demonstrated good test-retest reliability (ICCs for composite scores between .59 and .93), criterion validity (r s for composite scores between .66 and .87), and convergent/discriminant validity. Factorial validity and measurement invariance were demonstrated. These results compared favorably with those reported for the original interviewer version of the ASI and the English version of the ASI-MV.

Keywords

Addiction Severity Index; Spanish; assessment; validity

1.0 Introduction

The Addiction Severity Index (ASI) was created to assess current and lifetime problem severity in seven areas: alcohol use, drug use, employment, medical disorders, psychiatric disturbances, family/social relations, and legal problems (McLellan et al., 1985; McLellan et al., 1992). Over the years, the reliable and valid English version of the ASI (e.g., McLellan et al., 1985, 2004, 2006; Alterman et al., 1994; Cacciola et al., 1996) has become the standard in various research and clinical settings. This article describes the development and psychometric testing of a Spanish adaptation of a self-administered multimedia version of the ASI, the Addiction Severity Index-Multimedia Version or ASI-MV (Butler et al, 2001).

[†]Corresponding author. Contact Information: Inflexxion, Inc., 320 Needham St., Suite 100, Newton, MA 02464. Telephone 617-332-6028. Fax: 603-672-3162. E-mail address: sfbutler@inflexxion.com.

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.

The Hispanic population in the United States is expected to triple between 2000 and 2050 (U.S. Census Bureau, 2000). In 2003, 11% of Americans spoke Spanish at home, and of those, 28% spoke insufficient English (U.S. Census Bureau, 2003). A significant proportion of Hispanics (33.5%) indicate lifetime use of illegal drugs (SAMHSA, 2004). Thus, it is important that standard assessments, like the ASI and the ASI-MV, be adapted to Spanish to help remove cultural and language barriers to accessing treatment (Terrell, 1993). This article describes a Spanish adaptation of the ASI-MV, developed and validated for three Spanish populations in the United States.

The ASI scale and manual have been translated into Spanish (<http://www.densonline.org/DENShispASI.pdf>). Exhaustive searches of Medline and PsychINFO revealed no published accounts of efforts to evaluate this Spanish version of the ASI. This is critical given concerns about the effects of cultural misunderstandings on content validity of adapted psychological assessments (cf. Geisinger, 1994; Rogler, 1999). Despite this, there is widespread clinical and research use of the ASI (e.g., Robles et al., 2004) with Spanish-speaking populations. While clinicians must do their best with available assessments, use of unvalidated assessments is at odds with accepted standards for testing as put forth in the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1985).

Lack of a validated Spanish ASI, is not the only barrier to achieving a reliable and valid assessment of problem severity in Spanish-speaking substance abuse treatment consumers. The ASI assessment (in any language) is complex, requiring extensive and expensive training and monitoring of inter-rater reliability (Mäkelä, 2004). Given an average yearly turnover rate of 50% for substance abuse counselors (Carise et al., 2003), ongoing, adequate training is almost impossible (cf. Butler et al, 2001). Many authors, including McLellan and colleagues (2006), have bemoaned issues of poor inter-rater reliability with the ASI. Furthermore, even well-trained interviewers experience rater drift over time, requiring regular and expensive re-training.

Despite such problems, the ASI is increasingly mandated of clinical agencies (McLellan et al., 1992) and is arguably the standard assessment in the substance abuse field. Clinics with diminishing funds are required to both cut costs *and* provide extensive and standardized assessments, raising serious questions about the quality and relevance of ASI data being collected, issues only magnified when a Spanish-language ASI is considered. Clinicians treating Spanish-speaking clients have told us that they simply translate the ASI “on the fly,” thus further reducing standardization of assessment.

To address these problems for English-speaking clients, we developed the ASI-MV. The ASI-MV eliminates the need for interviewer training and standardizes patient scoring. The ASI-MV is reliable, valid, and has high user acceptance (Butler et al., 2001) and generates composite and severity scores for each ASI domain (McLellan et al, 1985). Butler and colleagues (1998) developed algorithms that mathematically predict severity ratings assigned by highly trained raters, so that this score can be provided without an interviewer. The ASI-MV has been well received and is used for treatment planning, program evaluation, and research; at this writing, the ASI-MV is administered in more than 400 facilities in 48 states.

This study describes the development and validation of a Spanish version of the ASI-MV (S-ASI-MV). Using state-of-the-art methods for adapting psychological measures to other cultures and languages (e.g., Geisinger, 1994; Rogler, 1999), the script of the English ASI-MV was adapted to the three largest Spanish-speaking populations in the U.S.: Mexican-American (66.%), Cuban-American (3.7%), and Puerto Rican (8.6%) (Ramirez et al., 2002). Note that

the Puerto Rican figures do not include the population of Puerto Rico, all of whom are U.S. citizens. Furthermore, 14.3% of the U.S. Spanish-speaking population is from Central and South America, which represent numerous culturally and linguistically distinct populations. A comprehensive test of reliability and validity was then conducted, along with an assessment of factorial structure and measurement invariance between the Spanish and English versions of the ASI-MV.

2.0 Methods

2.1 Adaptation Procedures

Adaptation of the ASI-MV to Spanish followed guidelines proposed by Geisinger (1994), who suggested that adaptation requires: (1) translation of the assessment by a person who is fluent in both languages, is knowledgeable about both cultures, and is an expert in the content to be measured; (2) individuals meeting these same requirements should carefully review the adapted measure; (3) feedback from the reviewers is integrated into the translation; (4) pilot testing; and (5) validation of the adapted instrument. Given that there are multiple Spanish-speaking cultures in the U.S., it was necessary to include translators representing three primary U.S. Spanish-speaking cultures: Mexican-American, Cuban-American, and Puerto Rican.

All translators were “balanced bilingual” professional substance abuse counselors (i.e., equally proficient in both languages). Furthermore, all translators/counselors resided in the U.S. and were self-identified as members of their respective cultural groups. Degree of bilingualism was assessed using a modified form of the Language History Questionnaire, which has shown high split-half reliability (.85), as well as significant correlations with key constructs (e.g., age of acquisition, years of learning) and discriminant validity (e.g., low, moderate and high proficiency; Li et al., 2006). Counselor applicants with a fluency score of “very good” or higher were interviewed by author, Dr. Redondo, who is fluent in English and Spanish, to confirm each candidate’s bilingual ability as well as his or her experience with clients in the target subculture.

Three bilingual substance abuse counselors (average years treating non-Hispanic and Hispanic substance abuse clients averaged 23.3 years, range 13 to 35 years) from the three target cultural backgrounds translated the English ASI-MV script to Spanish and were asked to indicate how they would say the text to a client in Spanish, focusing on the *meaning* implied rather than a word-for-word translation. The three translations were reviewed Dr. Redondo, who looked for commonalities and discrepancies, which were noted but *not resolved at this stage*. This effort resulted in Script 1.

Script 1 was evaluated by a panel of reviewers per Geisinger (1994). Another three bilingual substance abuse counselors (average years treating substance abuse clients = 17.7 years, range 16 to 22 years; average years treating Hispanic clients = 17.0 years, range 14 to 22 years) who met the same criteria as the translator/counselors were recruited and interviewed to verify fluency. After this second group reviewed the script and suggested changes, a consensus wording was achieved which resulted in Script 2.

Script 2 was pilot-tested, using “read through” procedures. Each question was read to fifteen substance abuse clients (mean age of 35.7, SD = 12.7, range 19 to 60, 40% women) and problems in readability or usability were noted. Finally, an independent, professional translator back-translated the Spanish script to check for errors. These steps resulted in a final script that was used to create the computer program to be psychometrically evaluated.

2.2 Participants

Adult consumers of residential and outpatient services for substance abuse were recruited. Inclusion criteria included: primary diagnosis of substance abuse or dependence, in treatment for at least three days, deemed stable enough to complete study procedures, and willingness to sign informed consent. Stabilization of the client was a subjective decision made by the treatment staff. An Institutional Review Board (IRB) reviewed and approved the consent forms and procedures for this study, including pilot testing. All subjects completed an informed consent form and were assured that the information obtained from the study would remain confidential and would not be a part of their clinical record nor affect their treatment in any way. Subjects were compensated for participation, and all were native Spanish speakers.

2.3 Treatment Facilities

The studies were conducted at eight substance abuse treatment centers in California, New Mexico, Puerto Rico and Florida. Sites were selected to achieve a multi-state and multicultural sample of clients from a variety of treatment settings. Two sites were located in California, both of which were day treatment centers; two sites in Puerto Rico, a residential facility and an outpatient center; three sites in Florida, two outpatient and one residential facility; and one outpatient site in New Mexico.

2.4 Measures

2.4.1 Interview Version of the Spanish ASI and Training of Interviewers—Criterion validity of the ASI-MV involved a test against the “traditional” interviewer version. However, determining this criterion was more complicated for the S-ASI-MV as there is no universally accepted interview version of the Spanish ASI. Given problems associated with the existing translated version, we were reluctant to use that as the “gold standard” against which to compare the S-ASI-MV. Instead, the S-ASI-MV script was retooled as an interview, so that questions and response options were printed for the interviewer to refer to during the interview.

Eight staff members were referred by the participating sites to undergo training to administer the face-to-face Spanish ASI interview. Two-day training, conducted by Dr. Redondo along with an ASI expert on the research team, covered the goals and objectives of the ASI, reviewed each question, and incorporated skill building exercises such as role playing, and scoring of video vignettes of actors.

During development of the English ASI-MV, ASI interviewers were trained using standard training offered by DeltaMetrics, which does not include demonstration of inter-rater reliability for the severity ratings. The likely unreliability of interviewers in that study resulted in some difficulty interpreting psychometric results for the English ASI-MV (see Butler et al., 2001). Given the importance of the interviewer results in establishing criterion validity for the S-ASI-MV, in the present study emphasis was put on achieving inter-rater agreement for severity ratings, and a checklist was created for interviewers to refer to when generating severity ratings. To test inter-rater reliability, two Spanish language video tapes of sample interviews were created, and trainees assigned severity ratings. Excellent inter-rater reliability was obtained among the trainees across all domains (IntraClass Correlations (ICC) of .99 for tape 1 and .91 for tape 2). Following the training, interviewers conducted all the interviews at their treatment site independently.

2.4.2 ASI Domain Comparison Measures—Selecting contrasting measures against which to test discriminant validity presents a greater challenge for a Spanish language assessment than for an English assessment (Rogers et al., 1995). Many translated measures are not formally validated, and even those with validation studies tend to be limited to the culture (s) tested, raising concerns of culturally appropriate content validity (Rogler, 1999). Finding

validated contrast measures for each of the ASI domains with the three targeted cultures proved impossible.

A search was conducted for Spanish language measures with at least some validity data reflecting the content areas of the ASI domains (with the exception of legal, as described below) and any available psychometric data were reviewed and the culture within which the measure had been tested noted. Selected measures were those with the best evidence of reliability and validity in any Spanish-speaking population, and when possible, previously tested on populations closest to our target populations. While we were able to find measures for each domain, the “match” was not perfect. Thus, bilingual researchers were instructed to help clients with interpretations of idiomatic expressions with which they had difficulty. Finally when possible, Spanish versions of measures used in the original validation of the ASI or ASI-MV were selected. Thus, (1) for the ASI medical domain: the Spanish translation of the general health subscale of the SF-36, which has been tested on Spanish speakers (Ware et al., 1993; Arocho et al., 1998 a,b) and measures four physical health concepts. (2) For the ASI employment domain: the work subscale of the 54-item, self report Social Adjustment Scale-Self Report (SAS-SR; Weissman & Bothwell, 1976) measures instrumental or expressive role performance for work functioning. While reliability and validity of the SAS-SR are adequate, the Spanish adaptation has not yet been validated. (3) For the ASI alcohol domain: the Michigan Alcoholism Screening Test (MAST; Selzer, 1971) and the alcohol problems subscale of the Personality Assessment Inventory (PAI; Morey, 1991) were used. The 10-item MAST assesses seriousness of alcohol abuse, and a translated version has been tested on a sample of Cuban medical patients (Gonzalez Menendez & Donaire Calabuch, 1989) and in Chile (Carmona, 1990). The well-validated alcohol problems subscale of the self-administered Personality Assessment Inventory (PAI; Morey, 1991) focuses directly on the problematic consequences of alcohol use and features of alcohol dependence. (4) For the ASI drug domain: the drug problems subscale of the PAI focuses on problematic consequences of drug use and features of drug dependence. Rogers et al. (1995) examined a translated version of the PAI with a population of Mexican-American substance abuse clients and found good test-retest reliability and modest to good internal consistency; (5) The ASI legal domain presents difficulties with respect to an independent measure assessing one’s legal problem severity. McLellan et al. (1985) reported finding no appropriate validating instrument for the legal problem area. Instead, they used reports of the total number of arrests and total months incarcerated. Following this lead, we collected self reports of the number of times the respondent had been arrested and total months incarcerated. (6) For the ASI family/social domain, the Family Unit Subscale of the Social Adjustment Scale Self Report (SAS-SR, Weissman & Bothwell, 1976) was used. Finally, (7) for the ASI psychiatric domain, the Hopkins Symptom Checklist-90 Item (HSCL-90; Derogatis et al., 1974) assesses a wide array of psychiatric symptoms. A Spanish adaptation of the SCL-90 has been tested in Argentina with physically ill and healthy community residents (Bonicatto et al., 1997) and with a population of university students in Spain (Henry et al. 1994).

2.5 Procedures

The psychometric evaluation of the S-ASI-MV involved examination of test-retest reliability, criterion validity, and discriminant validity. Factorial validity of the S-ASI-MV was examined using first order confirmatory factor analytic (CFA) techniques. Finally, measurement invariance of the S-ASI-MV with the English version of the ASI-MV was examined.

2.5.1 Procedures to Examine Test-Retest Reliability—Seventy-seven participants completed the S-ASI-MV at two time points, separated by three to five days (see McLellan et al., 1985). After a brief introduction to the computer program, all clients completed the S-ASI-

MV without assistance. Three to five days later, a return visit was scheduled to complete the S-ASI-MV a second time.

2.5.2 Procedures to Examine Criterion Validity—Criterion validity was evaluated by comparing the S-ASI-MV to the interviewer-administered version of the adapted S-ASI-MV script. Eighty-three Spanish-speaking participants were administered the S-ASI-MV and the adapted interview version ASI three to five days apart. Order of administration was counterbalanced to minimize any order effect.

2.5.4 Procedures to Examine Convergent/Discriminant Validity—All participants (N = 185) completed the S-ASI-MV followed by the comparison measures. Some participants also served in the test-retest study or the criterion validity studies. Since those who served in these two studies were also administered the battery of comparison measures, their data were used to examine discriminant validity. All 185 participants were included in the factorial validity and invariance analyses.

2.5.5 Measurement Invariance Comparison Databases—To determine model invariance across languages, comparisons were made between the Spanish participants in this study and respondents to the English version of the ASI-MV for the entire assessment and each ASI domain separately. We compared the fit of the restricted model, in which factor loadings are kept constant across the language groups, with the fit of the unrestricted model where factor loadings are allowed to vary across groups. Traditional tests of measurement invariance involve calculating the difference in chi-square ($\Delta\chi^2$), which if significant, suggests non-model invariance (Byrne, 1994). This method, however, has been criticized due to the sensitivity of the χ^2 to sample size (e.g., Brannick, 1995; Cheung & Rensvold, 2002). Cheung and Rensvold (2002) have shown that the difference in the Comparative Fit Index (ΔCFI) is not prone to these limitations, and that a ΔCFI of more than .01, in favor of the unrestricted model, suggests non-model invariance (that is, ΔCFI less than .01 supports model invariance). Measurement invariance of the S-ASI-MV was tested against an existing database of 986 English ASI-MV administrations drawn from four studies conducted by our research team from 1997 to 2003.

3.0 Results

3.1 Client Participants' Characteristics

Demographic data were obtained from answers to the S-ASI-MV. Characteristics of the overall sample and by study are summarized in Table 1. Note the overall sample was comprised mostly of men (82.7%), all of whom were Spanish-speakers from Cuban (32.4%), Mexican-American (37.8%), and Puerto Rican (26.5%) backgrounds (with 3.2% indicating unknown). Most participants cited cocaine (37.8%) as their primary substance of abuse, followed by alcohol (23.2%) and heroin (16.2%). Finally, mean education was 10.5 years (SD = 3.0) with more than a quarter of the sample (27%) having less than 9th grade education and more than half (56.2%) having less than a high school education.

3.2 Test-Retest Reliability

Test-retest reliability correlations are presented in Table 2 for the composite scores and for the computer-generated estimation of the severity rating. Of the 77 clients in this study, 75 (97.4%) completed both the Time 1 and Time 2 testing session. Both Pearson correlations and ICCs are presented. Test-retest for the composite scores are good, ranging from a low ICC of .59 (.60 Pearson *r*) for the family domain, to a high of .93 (ICC and Pearson) for the employment domain. For the severity ratings, the test-retest reliability coefficients were also good, ranging from a low of .69 (ICC and Pearson) for the employment domain to a high of .89 (ICC and Pearson) for the psychiatric domain.

3.3 Criterion Validity

Criterion validity contrasts the measure to be validated against some criterion or “gold standard” (Anastasi, 1976; Nunnally, 1978). For the S-ASI-MV, the “gold standard” was the interviewer-administered ASI. Of the 83 participants in the study, 97.6% (N = 81) completed both administrations. The correlations obtained between composite scores for the S-ASI-MV and the interviewer-administered version were positive, highly significant ($p < .001$), and generally of reasonable magnitude. As can be seen in Table 2, the correlations for the composite scores between the S-ASI-MV and the interviewer-administered version ranged from an r of .66 for the drug domain to a high of .87 for the alcohol domain.

To determine inter-rater reliability, one must view the S-ASI-MV as an “interviewer” whose inter-rater reliability must be established with live interviewers who have undergone the standard ASI interviewer training program. From this perspective, the ICC, which is typically used as an index of inter-rater reliability to examine the correspondence of the ratings of two or more raters, may be the most appropriate statistic (McGraw & Wong, 1996). Interpretations of the magnitude of ICCs generally assume that values greater than .80 represent perfect agreement, .61 to .80 is substantial, .41 to .60 is moderate, and .21 to .40 is fair reliability (Landis & Koch, 1977).

Examination of the ICCs demonstrates that reliability of the S-ASI-MV composite scores and severity ratings is good. For composite scores, the lowest ICC is for the drug domain (ICC = .64, in the “substantial” agreement range) with ICCs for both the employment and alcohol domains (ICC = .93 and .83, respectively, both of which are in the “excellent” agreement range). Similarly, for severity ratings, the lowest is for the employment domain (ICC = .60, in the “moderate” agreement range), with all other domains in the “substantial” agreement range.

3.4 Construct Validity (Convergent/Discriminant Validity)

To demonstrate construct validity, we replicated the convergent/discriminant validity test for the ASI (McLellan et al. 1985) and the English ASI-MV (Butler et al., 2001). This assessment of validity requires three conditions be met (Campbell and Fiske, 1959). First, each ASI-MV domain score (composite or severity) should be correlated in the appropriate direction with its designated comparison test. Second, each domain score should be more highly correlated with its designated comparison than with any of the other tests. Finally, a comparison test should be more highly correlated with its paired ASI domain score than with any other ASI domain score.

Reasonable variation was obtained for the comparison measures. Sample means and standard deviations for the comparison measures are as follows: general health scale of the SF-36 (transformed score): 65.2 (24.9); SAS-SR work scale: 1.9 (1.2); MAST total score: 11.2 (9.0); PAI alcohol score: 13.4 (8.2); PAI drug score: 18.9 (8.2); number of lifetime arrests: 6.1 (13.3); lifetime number of months incarcerated: 29.3 (50.8); SAS-SR family unit subscale: 2.0 (1.0); SCL-90 average score: 1.8 (0.7).

The correlation matrix for the composite scores is presented in Table 3 for the S-ASI-MV and the standard interviewer version. The boxed values are the correlations among the comparison tests with their paired ASI domain score. There was general evidence of good discriminant validity for the S-ASI-MV and, in most cases, the S-ASI-MV composite score correlated higher with the comparison measures (on-diagonal correlations) than with the other comparison measures (off-diagonal correlations). The family/social composite score correlated higher with the SCL-90 whereas the employment composite score had a correlation of only .13 with its comparison measure and the legal composite score correlation was nearly zero, patterns previously observed by both McLellan et al. (1985) and Butler et al. (2001) and may reflect a

unique focus of the ASI version of employment and legal problem severity scales. Nevertheless, the evidence for discriminant validity remains relatively strong. The average on-diagonal correlation (.36) was significantly higher than the average off-diagonal correlation (.13) for the S-ASI-MV ($t = 2.59$; $df = 8.6$ (equal variances not assumed), $p = .03$). Interestingly, the convergent/discriminant validity of the composite scores for the interviewer-administered ASI was less evident. Although the basic pattern of relationships was similar for both versions of the Spanish ASI, the magnitude of the correlations with the comparison measures was generally lower for the interviewer-administered Spanish ASI than for the S-ASI-MV. Furthermore, the average on-diagonal (.28) and off-diagonal (.12) correlations for the interviewer version were not significantly different ($t = 2.03$; $df = 8.7$, $p = .07$). This finding replicates previous results on the English ASI-MV (Butler et al., 2001) but differs from results observed in the initial test of the original ASI (McLellan et al., 1985).

Similar results were obtained for severity ratings (Table 4). There was general support for the discriminant validity of the S-ASI-MV severity ratings, with the average on-diagonal correlation (.42) significantly higher than the average off-diagonal correlation (.15; $t = 2.91$; $df = 8.7$, $p = .02$). Furthermore, average correlations for interviewer-generated severity ratings were not significantly different (on-diagonal mean = .29; off-diagonal mean = .12, $t = 1.41$; $df = 8.4$, $p = .19$).

3.5 Factorial Validity and Measurement Invariance of the S-ASI-MV

To examine the underlying factor structure of the S-ASI-MV, we conducted confirmatory factor analysis (CFA; Bentler and Stein, 1992). The S-ASI-MV's factorial validity model (first-order CFA model; Byrne, 2001) treats each ASI composite domain score as a factor (latent variable) that is predicted by that domain's composite items (observed variables), further stipulating that the S-ASI-MV domains (latent variables) are not independent (i.e., are likely to be correlated). This model yielded $\chi^2_{(188)} = 302.9$, $p < .001$. Although non-significant chi-squares are desirable, it is common for models to achieve significance, so generally an examination of fit indices is in order. The comparative fit index (CFI; Bentler, 1990) was acceptable (CFI=.91; generally values greater than .90 are considered acceptable). Examination of the root mean square error of approximation (RMSEA) statistic provides information about the covariance structure of the model and in this case was an acceptable .059 (values between .05 and .08 reflect a reasonably good fit, Byrne, 2001). These results confirm a reasonable factor structure for the S-ASI-MV.

Next, we examined the measurement invariance of the S-ASI-MV against an existing database of English ASI-MV administrations. Clients in the English database had an average age of 36.7 (SD = 9.0; range 18 to 68), with 38.9% women clients, 66.0% white clients, 21.6% African American, 5.9% Hispanic, 1.4% Native American, .4% Asian/Pacific Islander, and 4.4% unknown or missing. Following procedures suggested by Cheung & Rensvold (2002), an unconstrained model for the English and Spanish ASI-MV versions yielded $\chi^2_{(376)} = 1,224.4$, $p < .001$ with CFI = .894 and RMSEA = .044. Constrained models were run for the entire ASI (all domains constrained) as well as for each ASI domain separately. As shown in Table 5, the Δ CFI was less than .01 for all comparisons except for the employment domain. Thus, model invariance was observed for the overall ASI-MV, as well as for each domain except for employment, for which the model suggested non-model invariance.

3.5.1 Post Hoc Exploration of the ASI-MV and S-ASI-MV model fit—Given the poor fit observed above, we were interested to explore how respecification of the model might improve model fit. Having demonstrated generally good model invariance between the ASI-MV and the S-ASI-MV, these datasets were combined and the CFA was rerun. This yielded a $\chi^2_{(188)} = 981.1$, $p < .001$ with CFI = .90 and RMSEA = .060. Examination of standardized

estimates identified two variables as detracting from the model fit: for the legal domain, whether the client is currently awaiting trial, and for the family domain, serious conflicts in past 30 days with family members. Modification Indices (MI) for regression weights revealed a crossloading of clients' rating legal problems on the family latent variable (MI = 57.7). MIs for covariances revealed correlations between the error terms of all the alcohol and drug manifest variables. Given this evidence of misspecification and that relationships between these variables is logical, the model was respecified accordingly, which produced a better fitting model ($\chi^2_{(126)} = 449.1, p < .001$ with CFI = .96 and RMSEA = .047). While such respecification is interesting and suggests, perhaps, a connection between family and legal problems as well as alcohol and drug problems, such post hoc analyses must be considered exploratory only (Byrne, 2001).

4.0 Discussion

4.1 Summary of Findings

This study examined the reliability and validity of a Spanish adaptation of the ASI-MV. The S-ASI-MV was tested with native Spanish-speaking substance abuse clients representing Mexican-American, Cuban-American, and Puerto Rican populations. As with the English version, the S-ASI-MV demonstrated acceptable psychometric properties, including test-retest reliability criterion validity for both composite and severity scores.

There was also good evidence for convergent/discriminant validity. The results of this study are similar to those obtained for the English ASI-MV (Butler et al., 2001), and they further suggest that the S-ASI-MV appears at least as valid as the traditional, interviewer-administered ASI. In some cases, (i.e. drug and psychiatric domains) the S-ASI-MV achieved larger correlations with comparison measures than with the interviewer version, possibly representing a greater willingness of clients to disclose sensitive topics like drug use or psychiatric symptoms to a computer to an interviewer. Or, it may be that this pattern of results reflects some other method variance, such that self-reports (i.e., the comparison measures and S-ASI-MV) will correlate more highly with each other than with an interviewer-administered version. However, the high correspondence obtained between the S-ASI-MV and interview versions obtained in the criterion validity test (Table 2) might argue against a hypothesis that the apparently superior discriminant validity for the computer administration over the interview version is due only to this type of method variance.

Factorial validity was established with reasonably good model fit. These results reflect a superior model fit to those reported by Currie and colleagues (2004) for reasons that are unclear. One difference is the present examination of a Spanish language ASI. The superiority of the convergent/discriminant results for the ASI-MV over the interviewer ASI version may suggest something about computer-mediated administration of the ASI that results in a superior relationship between observed and unobserved ASI variables. Further work is warranted to replicate the current finding for the S-ASI-MV and to explicate the differences between these results and those reported by other investigators using different methods and languages.

In general, measurement invariance was observed in a comparison of the Spanish and English ASI-MV. Model invariance was observed for the S-ASI-MV both as a whole and when all domains were considered separately with the exception of the employment domain, which was not invariant. Indeed, other researchers have documented psychometric difficulties with the ASI employment domain (e.g., Alterman et al., 1998) which may have impacted the present results.

4.2 Acceptance of the S-ASI-MV

On-site researchers kept logs of incidents of rejection of the program or noticeable confusion when using it by any participant. There were no reports of rejection or confusion. Nor were such concerns expressed by the Spanish-speaking counselors. Indeed, spontaneous reports from several Spanish-speaking clients suggested they were pleased with the attention to their culture and language represented by the study and S-ASI-MV.

4.3 Limitations

Although we were careful to ensure successful adaptation of the S-ASI-MV to three, predominant Latino cultures within the U.S., the S-ASI-MV was not systematically tested with other Spanish-speaking cultures inside or outside the U.S. The goal was to achieve Spanish wording of questions and response options that would be generally understood by most Spanish speakers. Nevertheless, actual tests with other Spanish-speaking populations remain to be conducted. Caution should be used when administering the S-ASI-MV in Spain, Central and South America or with individuals from these populations within the U.S. (except for Mexican-Americans).

4.4 Clinical and Research Use of the S-ASI-MV

The S-ASI-MV shares many advantages of the original English ASI-MV, which has a high degree of acceptance in the real world of substance abuse treatment. Like its English predecessor, the S-ASI-MV is practical for use in clinical and research settings. Administering the S-ASI-MV requires little clinician time and produces computer-generated scores and reports in English.

Like the ASI-MV, the S-ASI-MV reduces the bias and inconsistency that can be introduced by undertrained interviewers by allowing for uniformity across administrations. The ASI-MV has proven cost-effective for treatment centers (Budman et al., 2008). We believe the same would be true for the Spanish version.

Limitations regarding the clinical use of the S-ASI-MV are similar to those for the ASI-MV and the interview version of the ASI. A client's current mental status must be appropriate for an extended assessment session, whether in person or on a computer. Disorganized, highly agitated, intoxicated clients or clients actively in withdrawal, of limited cognitive capacity or otherwise inappropriate to be interviewed should not be administered the S-ASI-MV. The S-ASI-MV, like the ASI-MV and the interview version of the ASI, is not intended as an assessment for adolescents younger than 18 (National Institute on Drug Abuse, 1993).

During the study, no volunteer participant rejected the computer administration. It should be highlighted that more than half (56%) of the sample had less than a high school education and more than a quarter (27%) reported education levels below 9th grade. Nevertheless, it is possible under ordinary clinical conditions, that some individuals might reject the computer based on a pre-existing bias against computers. There are about 80,000 ASI-MV administrations per year (Budman et al., 2008) and such events have been extremely rare although we have not systematically collected data on rejections in clinical settings. A few anecdotal instances involved individuals who were cognitively impaired. A positive and matter-of-fact attitude on the part of the staff person who introduces the S-ASI-MV should minimize rejection rates by clients.

4.5 Conclusion

This study found the S-ASI-MV to be a reliable and valid instrument for use with several Spanish-speaking populations in the U.S. The S-ASI-MV performs psychometrically as well or better than an interviewer version and has the advantage of presenting a consistent and

carefully adapted version of the ASI to this particular population. Like the ASI-MV, the S-ASI-MV should be as acceptable for most clinical and research purposes as the standard interview version of the ASI. Finally, the S-ASI-MV goes some distance toward addressing a barrier to treatment for Spanish speakers, a large and growing component of the U.S. population.

References

- Alterman AI, Brown LS, Zaballero A, McKay JR. Interviewer severity ratings and composite scores of the ASI: a further look. *Drug and alcohol dependence* 1994;34:201–209. [PubMed: 8033757]
- Alterman AI, McDermott PA, Cook TG, Metzger D, Rutherford MJ, Cacciola JS, Brown LS. New scales to assess change in the Addiction Severity Index for the opioid, cocaine, and alcohol dependent. *Psychology of Addictive Behavior* 1998;12:233–246.
- American Educational Research Association. Standards for educational and psychological testing. Washington, DC: American Psychological Association; 1985. American Psychological Association, & National Council on Measurement in Education.
- Anastasi, A. *Psychological testing*. Vol. 4. New York: Macmillan; 1976.
- Arocho R, McMillan CA. Discriminant and criterion validation of the U.S.-Spanish version of the SF-36 health survey in a Cuban-American population with benign prostatic hyperplasia. *Medical Care* 1998a; 36:766–772. [PubMed: 9596068]
- Arocho R, McMillan CA, Sutton-Wallace P. Construct validation of the USA-Spanish version of the SF-36 health survey in a Cuban-American population with benign prostatic hyperplasia. *Quality of Life Research* 1998b;7:121–126. [PubMed: 9523493]
- Bentler PM. Comparative fit indices in structural models. *Psychological Bulletin* 1990;107:238–246. [PubMed: 2320703]
- Bentler PM, Stein JA. Structural equation models in medical research. *Statistical Methods in Medical Research* 1992;1:159–81. [PubMed: 1341656]
- Bonicatto S, Dew MA, Soria JJ, Seghezze ME. Validity and reliability of Symptom Checklist 90 (SCL-90) in an Argentine population sample. *Social Psychiatry & Psychiatric Epidemiology* 1997;32:332–338. [PubMed: 9299927]
- Brannick MT. Critical comments on applying covariance structure modeling. *Journal of Organizational Behavior* 1995:201–213.
- Budman, SH.; Butler, SF.; Villapiano, AJ. The Addiction Severity Index Multimedia Version (ASI-MV): An innovative computerized assessment tool for substance abuse treatment. In: O'Donohue, WT.; Cummings, NA., editors. *Evidence Based Adjunctive Treatments*. Burlington, MA: Academic Press/Elsevier; 2008. p. 141-158.
- Butler SF, Newman FL, Cacciola JS, Frank A, Budman SH, McLellan AT, Ford S, Blaine J, Gastfriend D, Moras K, Salloum IM. Predicting ASI interviewer severity ratings for a computer-administered Addiction Severity Index. *Psychological Assessment* 1998;10:399–407.
- Butler SF, Budman SH, Goldman RJ, Beckley KE, Trottier D, Newman FL, Cacciola JS. Initial validation of a computer-administered Addiction Severity Index: the ASI MV. *Psychology of Addictive Behaviors* 2001;15:4–12. [PubMed: 11255937]
- Byrne, BM. *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Lawrence Erlbaum; Mahwah, New Jersey: 2001.
- Byrne BM. Testing for the factorial validity, replication, and invariance of a measuring instrument: A paradigmatic application based on the Maslach Burnout Inventory. *Multivariate Behavioral Research* 1994;29:289–311.
- Cacciola J, Rutherford MJ, Alterman AI, McKay JR, Snider EC. Personality disorders and treatment outcome in methadone maintenance patients. *Journal of Nervous and Mental Disease* 1996;184:234–239. [PubMed: 8604033]
- Campbell DT, Fiske D. Convergent and discriminant validity by the multi-trait multi-method matrix. *Psychological Bulletin* 1959;56:81–105. [PubMed: 13634291]
- Carise, D.; McLellan, AT.; Gifford, LS. Developing a national addiction treatment information system. 2003 [12/2/03]. <http://www.densonline.org/DENSArt1update040101.pdf>

- Carmona A. El test de seleccion de bebedores problema (TBSP): Relacion con edad y con antecedentes hereditarios de alcoholismo/ Michigan Alcohol Screening Test (MAST): Relationship with age and family history of alcoholism. *Revista de Psiquiatria Clinica* 1990;27:9–16.
- Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling* 2002;9:233–255.
- Currie SR, el-Guebaly N, Coulson R, Hodgins DC, Mansley C. Factor validation of the Addiction Severity Index scale structure in persons with concurrent disorders. *Psychological Assessment* 2004;16:326–329. [PubMed: 15456388]
- Derogatis L, Lipman R, Rickels K. The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral Sciences* 1974;19:1–16.
- Geisinger KF. Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessments* 1994;6:304–312.
- Gonzalez Menendez R, Donaire Calabuch I. Los cuestionarios “CAGE” y “MAST” abreviado en un servicio de medicina interna: Sensibilidad y especificidad segun las fuentes do informacion directa o familiar/ CAGE and MAST abbreviated questionnaires in an internal medicine service: Sensitivity and specificity according to direct or family information sources. *Revista del Hospital Psiquiatrico de La Habana* 1989;30:197–207.
- Henry M, Morera A, Preciado J, Frugoni A, Gracia R, Morales CR. Validez concurrente de la escala de depresion del SCL-90-R y la S.D.S. de Zung/ Concurrent validity of SCL-90-R depression scale and Zung’s S.D.S. *Anales de Psiquiatria* 1994;10:235–237.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–174. [PubMed: 843571]
- Li P, Sepanski S, Zhao X. Language history questionnaire: A Web-based interface for bilingual research. *Behavior Research Methods* 2006;38:202–210. [PubMed: 16956095]
- McGraw KO, Wong SP. Forming inferences about some Intraclass Correlation Coefficients. *Psychological Methods* 1996;1:30–46.
- Mäkelä K. Studies of the reliability and validity of the Addiction Severity Index. *Addiction* 2004;99:398–410. [PubMed: 15049734]
- McLellan AT, Alterman AI, Cacciola J, Metzger D, O’Brien CP. A new measure of substance abuse treatment: Initial studies of the treatment services review. *Journal of Nervous and Mental Disease* 1992;180:101–110. [PubMed: 1737971]
- McLellan AT, Kushner H, Metzger D, Peters R, Smith L, Grissom G, Pettinati H, Argeriou M. The Fifth Edition of the Addiction Severity Index. *Journal of Substance Abuse Treatment* 1992;9:199–213. [PubMed: 1334156]
- McLellan AT, Cassiola JS, Alterman AI. The ASI is a still developing instrument: Response to Mäkelä. *Addiction* 2004;99:411–412.
- McLellan AT, Cacciola JC, Alterman AI, Samuel H, Rikoon SH, Carise D. The Addiction Severity Index at 25: Origins, contributions and transitions. *The American Journal on Addictions* 2006;15:113–124.
- McLellan AT, Luborsky L, Cacciola J, Griffith JE, Evans F, Barr HL, O’Brien CP. New data from the Addiction Severity Index: Reliability and validity in three centers. *Journal of Nervous and Mental Disease* 1985;173:412–423. [PubMed: 4009158]
- Morey, L. *Personality Assessment Inventory: Professional Manual*. Psychological Assessment Resources; Tampa, FL: 1991.
- National Institute on Drug Abuse (NIDA). *Assessing client needs using the ASI: A handbook for program administrators*. NIH pub No 93-3619. Rockville, MD: 1993.
- Nunnally, J. *Psychometric theory*. Vol. 2. New York: McGraw-Hill; 1978.
- Ramirez, RR.; de la Cruz, GP. *Current Population Reports*. U.S. Census Bureau; Washington DC: 2002. The Hispanic population in the United States: March 2002; p. P20-545.
- Robles RR, Reyes JC, Colón HM, Sahai H, Marrero CA, Matos TD, Calderón JM, Shepard EW. Effects of combined counseling and case management to reduce HIV risk behaviors among Hispanic drug injectors in Puerto Rico: A randomized controlled study. *Journal of Substance Abuse Treatment* 2004;27(2):145–152. [PubMed: 15450647]

- Rogers R, Flores J, Ustad K, Sewell KW. Initial validation of the Personality Assessment Inventory-Spanish version with clients from Mexican American communities. *Journal of Personality Assessment* 1995;64:340–348. [PubMed: 7722858]
- Rogler LH. Methodological sources of cultural insensitivity in mental health research. *American Psychologist* 1999;54:424–433. [PubMed: 10392472]
- SAMHSA. Office of Applied Studies, National Survey on Drug Use and Health, 2002, 2003, 2004.
- Selzer MD. The Michigan Alcoholism Screening Test: The quest for a new diagnostic instrument. *American Journal of Psychiatry* 1971;127:1653–1658. [PubMed: 5565851]
- Terrell MD. Ethnocultural factors and substance abuse: Toward culturally sensitive treatment models. *Psychology of Addictive Behaviors* 1993;7:162–167.
- U.S. Census Bureau. Summary File 1 (SF1) 100 Percent Data, generated by Julie Brevard using data extraction system. 2000 [January 15, 2005].
<http://factfinder.census.gov/servlet/DatasetMainPageServlet>
- U.S. Census Bureau. Selected social characteristics (table) 2003. *American Community Survey* 2003. 2003
- Ware, JD.; Snow, KK.; Kosinski, M.; Gandek, B. SF-36 Health Survey: Manual and Interpretation Guide. The Health Institute, New England Medical Center; Boston, MA: 1993.
- Weissman MM, Bothwell S. Assessment of social adjustment by patient self-report. *Archives of General Psychiatry* 1976;33:1111–1115. [PubMed: 962494]

Table 1
Demographic Characteristics of Participants in Overall Sample and by Study

Characteristic	Overall Sample Discriminant Validity N = 185 ¹	Test-Retest Reliability N= 77	Criterion Validity N=83
Age (in years)			
Mean (SD)	39.8 (11.2)	38.4 (8.8)	40.1 (9.5)
Range	18 – 69	18 - 63	18 – 69
Gender			
Male	82.7%	81.8%	83.1%
Female	17.3%	18.2%	16.9%
Cultural Background			
Cuban	32.4%	33.8%	38.6%
Mexican-American	37.8%	29.9%	37.3%
Puerto Rican	26.5%	35.1%	18.1%
Unknown	3.2%	1.3%	6.0%
Marital Status			
Married/live with SO	2.2%	0%	2.4%
Single	34.6%	31.2%	42.2%
Divorced/separated	54.0%	59.9%	47.0%
Widowed	9.2%	9.1%	8.4%
Education (in years)			
Mean (SD)	10.5 (3.0)	10.8 (3.4)	10.4 (2.7)
Range	6 - 20	6 - 20	6 - 20
% < 9 th grade	27.0%	24.7%	28.9%
% < high school	56.2%	58.4%	51.8%
Court ordered treatment	34.1%	32.5%	34.9%
Currently on parole or probation	27.6%	31.2%	21.7%
Primary Substance of Abuse			
Alcohol Only	23.2%	15.6%	22.9%
Heroin	16.2%	18.2%	18.1%
Cocaine	37.8%	40.3%	34.9%
Cannabis	2.9%	2.6%	4.8%
Amphetamines	4.3%	7.8%	2.4%
Sedatives	1.4%	2.6%	1.2%
Did not indicate primary	13.0%	13.0%	15.7%

¹ Since studies all Ss participated in more than one “study,” the Ns do not add.

Table 2

Test-retest and Criterion Validity Results

	Test-Retest Statistics n=75				Criterion Validity Statistics n=81			
	Composite		Severity		Composite		Severity	
	ICC	Pearson	ICC	Pearson	ICC	Pearson	ICC	Pearson
Medical	.675	.681	.747	.748	.742	.743	.619	.619
Employment	.930	.931	.685	.686	.849	.850	.604	.605
Alcohol	.832	.841	.857	.857	.861	.866	.703	.706
Drug	.777	.778	.783	.783	.644	.656	.633	.676
Legal	.689	.689	.747	.748	.673	.673	.695	.698
Family	.589	.603	.812	.812	.694	.734	.677	.680
Psychiatric	.879	.884	.889	.890	.681	.684	.731	.734

Table 3
Convergent and Discriminant Validity Correlations between S-ASI-MV Composite Scores and Comparison Measures

or Interviewer	SF-36 Health Scale	SAS Work Subscale	MAST	PAI Alcohol Subscale	PAI Drug Subscale	Number of arrests	Months Incarcerated	SAS Family Unit Subscale	SCL-90
	-.418	.144	.221	.127	.172	.017	.160	.276	.413
	-.325	.117	.215	.023	.029	.134	.101	.330	.257
	-.216	.131	.054	.143	.027	.105	.177	.094	.083
	-.119	.181	.066	.045	.015	-.041	.194	.077	.071
	-.236	.042	.606	.606	.046	.015	-.005	-.013	.349
	-.103	.208	.489	.466	.010	-.004	-.182	-.035	.053
	-.023	.007	.084	.103	.559	.006	.067	.264	.202
	-.054	.036	.280	.183	.399	.219	-.047	.220	.279
	-.007	.073	.153	.069	.152	.028	-.057	.010	.151
	.056	.084	.314	.215	.136	.018	-.185	.034	.150
	-.051	.031	.187	.110	.206	-.074	-.069	.341	.373
	-.147	.162	.189	.150	.300	.060	-.099	.376	.368
	-.324	-.204	.352	.334	.291	.105	.105	.284	.642
	-.085	.193	.235	.210	.138	.189	-.067	.073	.406

ored.

r ASI domain's comparison measure are presented in bold type.

Table 4
 Convergent and Discriminant Validity Correlations between S-ASI-MV Severity Ratings and Comparison Measures

or Interviewer	SF-36 Health Scale	SAS Work Subscale	MAST	PAI Alcohol Subscale	PAI Drug Subscale	Number of arrests	Months Incarcerated	SAS Family Unit Subscale	SCL-90
	-.404	.151	.220	.148	.181	.005	.136	.248	.437
	-.353	.151	.172	-.024	.023	.093	.104	.231	.276
	-.130	.050	.065	.095	.254	.063	.120	.214	.317
	.066	-.376	-.019	.025	.238	.175	.012	.210	.067
	-.184	-.005	.702	.691	.121	.02	-.015	.000	.257
	-.132	.201	.704	.600	.187	.015	-.133	.030	.060
	-.013	-.001	.078	.045	.640	.138	.220	.233	.223
	-.143	-.116	.196	.088	.410	.298	.135	.272	.213
	.013	.001	.161	.060	.197	.101	.128	.018	.129
	.079	-.195	.239	.164	.200	.023	.035	.173	.097
	-.109	.132	.144	.188	.309	.002	-.075	.427	.554
	-.172	.020	.206	.116	.284	.020	-.005	.304	.391
	-.271	.187	.387	.376	.366	.116	.094	.333	.608
	-.019	.029	.349	.352	.265	.193	-.070	.066	.302

ored.

ASI domain are in bold type.
 Drug Alcohol Depend. Author manuscript; available in PMC 2010 January 1.

Table 5

Summary of Goodness-of-Fit Statistics for Invariance for Spanish & English ASI-MV

Model	χ^2	<i>df</i>	CFI	ΔCFI
Unconstrained Model	1,224.4	376	.894	—
Fully Constrained Model	1,310.5	391	.885	.009
Medical Domain Constrained	1,237.8	378	.892	.002
Employment Domain Constrained	1,634.1	379	.843	.051
Alcohol Domain Constrained	1,249.5	378	.891	.003
Drug Domain Constrained	1,234.0	378	.893	.001
Legal Domain Constrained	1,242.2	379	.892	.002
Family/Social Domain Constrained	1,229.1	378	.893	.001
Psychiatric Domain Constrained	1,236.0	378	.893	.001