An Item Response Theory (IRT) Analysis of the Short Inventory of Problems-Alcohol and Drugs (SIP-AD) among non-treatment seeking Men-Who-Have-Sex-With-Men: Evidence for a shortened 10-item SIP-AD

Brett T. Hagman, M.A.[a], **Alexis N. Kuerbis, Ph.D.**[a], **Jon Morgenstern, Ph.D.**[b], **Donald A. Bux, Ph.D.**[a], **Jeffrey T. Parsons, Ph.D.**[c], and **Bram E. Heidinger, B.A.**[a]

[a]Research Foundation for Mental Hygiene; Columbia Addiction Services and Psychotherapy Interventions Research; 1775 Broadway, Suite 1404; New York, NY 10019

[b]Research Foundation for Mental Hygiene & Columbia University College of Physicians & Surgeons; Harkness Pavilion, Rm 236; 180 Ft. Washington Ave., New York, NY 10032

[c]Hunter College and the Graduate Center of the City University of New York; Center for HIV/AIDS Educational Studies & Training; 695 Park Avenue, HN611C,; New York, NY 10065

Abstract

The Short Inventory of Problems-Alcohol and Drugs (SIP-AD) is a 15-item measure that assesses concurrently negative consequences associated with alcohol and illicit drug use. Current psychometric evaluation has been limited to classical test theory (CTT) statistics, and it has not been validated among non-treatment seeking men-who-have-sex-with-men (MSM). Methods from Item Response Theory (IRT) can improve upon CTT by providing an in-depth analysis of how each item performs across the underlying latent trait that it is purported to measure. The present study examined the psychometric properties of the SIP-AD using methods from both IRT and CTT among a non-treatment seeking MSM sample (N = 469). Participants were recruited from the New York City area and were asked to participate in a series of studies examining club drug use. Results indicated that five items on the SIP-AD demonstrated poor item misfit or significant differential item functioning (DIF) across race/ethnicity and HIV status. These five items were dropped and two-parameter IRT analyses were conducted on the remaining 10 items, which indicated a restricted range of item location parameters ($-.15$ to $-.99$) plotted at the lower end of the latent negative consequences severity continuum, and reasonably high discrimination parameters (1.30 to 2.22). Additional CTT statistics were compared between the original 15-item SIP-AD and the refined 10-item SIP-AD and suggest that the differences were negligible with the refined 10-item SIP-AD indicating a high degree

of reliability and validity. Findings suggest the SIP-AD can be shortened to 10 items and appears to be a non-biased reliable and valid measure among non-treatment seeking MSM.

Keywords

Item Response Theory; Reliability; Validity; Alcohol Use; Drug Use

1. Introduction

Alcohol and club drug use is prevalent among men-who-have-sex-with-men (MSM) and remains a significant public health concern (Parsons, Halkitis, & Bimbi, 2006; Vaudrey et al., 2007). Several studies have noted that MSM consume alcohol at comparable rates to those in the general population, but that problem-drinking MSM are less likely to abstain from drinking (Bux, 1996; Stall et al., 2001). Interestingly, MSM are more likely to experience more negative consequences from drinking compared to heterosexual men (Bergmark, 1999). MSM also have been shown to exhibit higher rates of club drug use (i.e., defined as the use of MDMA/Ecstasy, GHB, ketamine, crystal methamphetamine, or cocaine) in comparison to their heterosexual counterparts (Koblin et al., 2006). Estimates of illicit drug use among MSM, which include club drugs, are as high as 19%, and drug-related problems among this at-risk group have been estimated at 5.7% (Cochran, Ackerman, Mays, & Ross, 2004).

Because of these elevated rates of club drug and alcohol use among MSM and their relative acceptance within gay communities (Harawa et al., 2008), the risk for experiencing negative consequences of drug and alcohol use remains high. Two other subgroups of MSM, African-Americans and those who are HIV+, may experience amplified or a unique set of negative consequences to drinking and drug use compared to the wider group of MSM. Racism and discrimination have been identified as central to alcohol and drug use among African-Americans (Wright, 2001). African-Americans are known to suffer from racially based health disparities, including significantly higher incidence and prevalence rates of HIV (Torian, Makki, Menzies, Murrill, & Weisfuse, 2002; Wheeler, Lauby, Liu, Van Sluytman, & Murrill, 2008), which may intensify negative consequences of substance use. Heightened homophobia existing in African-American communities (Wright, 2001) may also influence the perception of negative consequences of drug and alcohol use. Drug use has been identified as being a potential key component in same-sex sexuality among African-Americans (Harawa et al., 2008), and it may play a key facilitating role in coping with a homosexual or bisexual identity.

HIV infection may also influence an individual's response to the negative consequences of alcohol and drug use. One study found that a majority of MSM with HIV reduced alcohol and drug use, as recreational drugs increased the unpleasant side effects of combination anti-viral therapies (Nollen et al., 2002). HIV positive individuals who are not on medication regimens may not have such deterrents and may experience negative consequences differently than their counterparts who are medicated. Another study found that HIV+ men with a current substance use disorder were more likely to report a higher number of depressive symptoms, overall distress, and a lower quality of life than those without a substance use disorder (Ferrando et al., 1998). This may indicate that the unique combination of HIV infection and a substance use disorder may influence one's perceived negative consequences of alcohol and other drug use.

Assessment and measurement of negative consequences from the effects of both alcohol and club drug use among all groups of MSM are critical to both understanding the unique treatment needs of MSM and preventing future substance abuse and dependence in the MSM community. Substance use assessment tools that have been thoroughly vetted through extensive

psychometric evaluation in multiple ways, as well as validated across subgroups of the MSM population, are necessary to obtain that understanding.

Over the past decade, researchers have developed several measures for assessing negative consequences associated with the effects of alcohol and other drug use. The Inventory of Drug Use Consequences (InDUC-2R; Miller, Tonigan, & Longabaugh, 1995; Tonigan and Miller, 2002) was developed to assess for negative consequences associated with both alcohol and other drug use. This 50-item scale measures consequences of alcohol and drug use across five subscales, each focusing on a distinct domain: interpersonal, intrapersonal, physical, impulse control, and social. In an effort to shorten the InDUC-2R and enhance its utility in clinical settings, Blanchard and colleagues (2003) conducted a psychometric analysis of the InDUC-2R among a community outpatient sample of patients with both alcohol and other drug use problems. Psychometric analyses examined the item to total scale correlations selecting 15 items with the highest correlations to represent the targeted construct. The shortened measure resulted in the development of the 15-item SIP-AD, which demonstrated good internal consistency, convergent and discriminant validity, and was sensitive to detect changes pre-post assessment across a 3-month interval. Since its inception, psychometric evaluation of the SIP-AD across specific sub-populations has been scant (Gillespie, Holt, & Blackwell, 2007). The extent to which the psychometric properties of the SIP-AD hold among non-treatment seeking MSM has yet to be investigated.

Existing psychometric evaluation of the SIP-AD has focused exclusively on classical test theory (CTT) statistics (e.g., concurrent validity). CTT statistics are associated with certain disadvantages, such as being item sample dependent, and they do not provide information on how an individual or specific sub-group performs on a single item (Kahler, Strong, Hayaki, Ramsey, & Brown, 2003). Therefore, the quality of a measure is limited because it does not provide an in-depth analysis of how well the measure maps the underlying latent trait that it is purported to measure (Embretson & Reise, 2000). Methods from item response theory (IRT) can improve upon CTT because IRT permits assessment of the ordering of severity of each specific item within a measure. By examining the extent to which an item is endorsed across varying levels of the underlying trait, such as severity of alcohol and other drug use, it provides information about which items are indicative of more intense levels of the underlying trait, in this case more severe levels of alcohol and other drug problems (Neal, Corbin, & Fromme, 2006). In addition, IRT methods can provide information on how well an item places an individual on a particular point in the underlying latent-trait continuum (Neal et al., 2006). Another advantage of IRT methods is its ability to examine each item's differential item functioning (DIF) across select demographic variables in order to understand the extent to which certain items may be biased (Embretson & Reise, 2000).

Given the identified gaps in the literature, the purpose of the present study was to use methods based on IRT to provide a more in-depth analysis of the psychometric properties of the SIP-AD among a non-treatment seeking MSM sample. More specifically, there were three specific aims: 1) to examine the differential item functioning (DIF) of each of the SIP-AD items across race/ethnicity and HIV status; 2) to apply two-parameter IRT logistic models to the SIP-AD estimating both item *location (severity)* and *discrimination (slope)* parameters for each item and assess model fit for each item; and 3) to examine the reliability and validity of the SIP-AD using CTT statistics based on recommendations from the IRT analysis.

2. Methods

2.1 Participant Selection

Participants were non-treatment seeking MSM (N = 469) recruited between June 2005 to June 2006 through community outreach in New York City. Participants were eligible for the study

if they met the following criteria: 1) were male between the ages of 18 to 65; 2) reported a minimum of 10 occasions of club drug use in the prior 90 days; 3) had sexual contact with another man in the past 3 months; 4) not currently in treatment for substance abuse in the past month; and 5) reported that club drug use was at least as significant for them as alcohol use.

2.2 Recruitment

Recruitment consisted of a combination of print and online advertising in a number of gay-related media outlets, and active recruitment using outreach workers making contacts within the community (see Grov, Bux, Parsons, & Morgenstern, *in press* for a full description of recruitment strategies). Recruitment messages were designed to target minimally stigmatized behavior, directed at MSM with any history of ecstasy (MDMA) use. This strategy was based on previous research suggesting that MDMA use is commonly reported among MSM (Klitzman, Greenberg, Pollack, & Dolezal, 2002; Klitzman, Pope, & Hudson, 2000), and would cast the widest possible net for the present study.

2.3 Measures

All data were collected via self-report. To enhance accuracy, all participants were told that their responses would remain confidential. Specific domains captured are described below:

2.3.1 Demographics and HIV status—Several questions collected basic demographic information such as current HIV status, age, race/ethnicity, education and employment status.

2.3.2 Alcohol and other drug use—Participants were asked to report the number of days in the past 90 for which they used each of the following 5 club drugs (i.e., cocaine, methamphetamine, ketamine, GHB, and ecstasy), and alcohol. In addition, the average number of standard drinks consumed per week was assessed.

2.3.3 Drug dependence & SIP-AD—Two measures of problem severity of alcohol and drug use were collected. First, symptoms of DSM-IV drug dependence were assessed. Participants were asked to report whether they had experienced each of these symptoms as a consequence of any club drug use in the prior 90 days (Cronbach's alpha = .82); thus, participants' responses did not necessarily imply that all endorsed symptoms were associated with the same drug. Second, specific alcohol and club drug-related consequences were assessed via the SIP-AD (Blanchard et al., 2003), a 15-item measure that assesses the incidence and severity of drug and alcohol related problems simultaneously. The SIP-AD assessed the number of times (0 = none to 3 = daily) each participant experienced 15 different negative consequences from their alcohol or club drug use during the prior 90-day period. The SIP-AD has established reliability and validity, and yields a severity score ranging from 0–45 (Blanchard et al., 2003). For the purposes of the present study, each SIP-AD item was dichotomized to reflect either presence or absence of each specific consequence with total scores ranging from 0 to 15. Dichotomization of the response options of the SIP-AD was necessary due to low endorsement ratings for consequences that occurr multiple times (e.g., 3 = daily), which can cause problems with the estimation of IRT parameter estimates (Neal et al., 2006).

2.3.4 Alcohol and other drug treatment history—Two separate questions assessed if each participant had ever received formal treatment for a drug or alcohol problem as well as if he had ever attended any self-help groups/treatment.

2.3.5 Data Collection Procedure—Participants completed all measures on a computer via the Audio Computer Assisted Self-Interview (ACASI). Studies have shown that ACASI increases the proportion of individuals admitting drug use (Tourangeau & Smith, 1996; Turner,

Ku, Rogers, Lindberg, & Pleck, 1998). An interviewer provided initial training on the use of the computer program and administered the first several questions with the participant to ensure that he understood the procedure. The interview typically required 1 hour, and participants were compensated $30 for completion.

2.4 Analytical Plan

For the present study, two-parameter logistic models were applied to the SIP-AD items, which provide estimates of both item *location* or difficulty parameters (with values that typically range between −2 to 2) and the item *discrimination* or slope parameters (with values that typically range between 0 to 2). Specifically, larger location parameters indicate that elevated values of the underlying latent-trait are necessary in order to endorse the item (Neal et al., 2006). In addition, larger discrimination parameters signify that the probability of endorsing an item increases more quickly as the purported underlying latent trait increases (Neal et al., 2006).

IRT analyses were conducted via a series of steps: First, we conducted a confirmatory factor analysis (CFA) on the dichotomized SIP-AD items specifying a single factor solution to ensure that IRT assumptions were met. Next, an item misfit analysis was conducted in order to determine the extent to which each item fit the specified two-parameter model. Then, differential item functioning (DIF) was examined across race-ethnicity (African-Americans vs. Non-African Americans) and self-reported HIV status (HIV positive vs. HIV negative) and a model was built by allowing items to vary across the location parameters. The final IRT analyses examined those remaining items that did not demonstrate significant DIF based on race-ethnicity, HIV status or demonstrate item misfit. All further psychometric analyses were examined on this final set. Lastly, a total information curve, which is estimated by all values from the location and discrimination parameters for each item and indicates the amount of information (i.e., the point on the latent-trait where the scale is most reliable) the scale provides across the underlying latent-trait continuum, was plotted for the refined SIP-AD. All IRT models were analyzed using Parscale 4.1 (Scientific Software International, 2003), which estimates each item parameter via a Bayesian expectation-maximization (EM) estimation procedure. For the present analyses, the criterion applied for convergence of the EM estimation procedure was 0.005.

Additional analyses focused on examining the psychometric properties of the SIP-AD using CTT statistics based on recommendations made for scale refinement from the IRT analyses based on the original SIP-AD scoring scheme. Internal consistency reliability was assessed by calculating Cronbach's Coefficient Alpha, and additional analyses examined the item to total scale correlations. With respect to examining the validity of the SIP-AD, correlations with which adverse consequences would be correlated (e.g., severity of drug dependence) with the SIP-AD were conducted, as well as to determine its factor structure, an exploratory principal components analysis was implemented on the SIP-AD. Further, in order to provide a comparison and justification for the refined SIP-AD, CTT statistics were calculated on the original 15-item SIP-AD.

3. Results

3.1 Demographics of Current Sample

The majority of the sample were either African-American (37.2%, n = 175), White (31%, n = 146) or Hispanic/Latino (19.1%, n = 90) and had a mean age of 38.38 (SD = 9.67). Approximately 48.8% (n = 140) of the sample were currently employed, and more than half had some college education (69.6%, n = 328). In addition, most participants self-identified themselves as gay/homosexual (73.5%, n = 346), HIV positive (61.4%, n = 289), and had never

sought formal treatment for alcohol or illicit drug use (61.8%, n = 291). Participants reported using cocaine, methamphetamine, ecstasy, ketamine and GHB on 20.04 (SD= 23.2), 11.7 (SD = 18.2), 9.84 (SD = 14.7), 2.62 (SD = 8.4), and 1.84 (SD = 8.3) days during the prior 90, respectively. Further, alcohol use among the sample was common with participants drinking approximately 33.5 (SD = 27.6) days in the last 90 and consuming, on average, 14.2 (SD = 17.9) standard drinks per week.

3.2 Checking Model Assumptions

Two primary assumptions of IRT are local independence, which posits that when respondent trait levels are controlled for, the items on the scale are independent from one another, and unidimensionality of the scale (Embretson & Reise, 2000; Neal et al., 2006). In order to ensure this, we conducted a confirmatory factor analysis (CFA) on the 15 SIP-AD items. Since the SIP-AD items were dichotomized, the CFA was conducted using tetrachoric correlations. A single factor was specified and all factor loadings were permitted to vary. Overall, model fit was good: Tucker-Lewis Index (TLI) = .993, comparative fit index (CFI) = .956, root mean square error of approximation (RMSEA) = .059 and the factor loadings for each of the 15 SIP-AD items ranged from .862 to 1.156. In sum, the SIP-AD appears to have a unidimensional structure associated with it and IRT assumptions were met.

3.3 Item Misfit Analysis

The item misfit analysis examined the extent to which each item fits the specified two-parameter model with statistically significant chi-square values indicating poor item misfit. Overall, the two-parameter model fit the data reasonably well, except for item #6 (While using, I have said harsh or cruel things to someone), X^2 (1, N = 469) = 31.56, p < .01, and item #5 (I have taken foolish risks when I have been using), X^2 (1, N = 469) = 13.12, p < .01.

3.4 DIF Analysis

For the purposes of this study, DIF analysis were examined between AA vs. non-AA MSM and HIV positive vs. HIV negative MSM. As noted by Neal and colleagues (2006), significant DIF on the location parameters is *uniform* and indicates relative differences between the two groups. Differences across race-ethnicity and HIV status with respect to differential item functioning were tested using chi-square analyses. The reference groups selected for these analyses were the AA sub-group and HIV positive sub-group in which the severity estimates were anchored with M = 0 and SD = 1. Table 1 displays the probability of endorsement for each item, location and discrimination parameters, and results from the DIF analysis between AA and non-AA MSM and HIV positive and HIV negative MSM. With respect to differences across race/ethnicity, significant DIF on the location parameters occurred for item #2 (Because of my use, I have lost weight or not eaten properly), X^2 (1, N = 469) = 3.51, p < .05, and item #12 (I have lost interest in activities and hobbies because of my use), X^2 (1, N = 469) = 3.59, p < .05. This indicates that at equivalent levels of alcohol and other drug use severity, AA MSM were more likely to endorse these two items than non-AA MSM. In addition, significant DIF between HIV status occurred for item #10 (My family has been hurt by my use), X^2 (1, N = 469) = 5.2, p < .02. This indicates that at equivalent levels of alcohol and other drug use severity, HIV positive MSM were more likely to endorse this item than HIV negative MSM.

3.5 Final IRT Model Analysis

For the final IRT analysis, on the basis of significance testing from both the DIF and item misfit analyses, we decided to drop items #2, #5, #6, #10, and #12. The final two-parameter model was conducted on the remaining 10 items. Results from the final IRT model are presented in table 2. The mean number of consequences endorsed for the refined 10-item SIP-AD among the sample was 6.71 (SD = 3.42). Overall, there was high degree of endorsement for each of

the individual SIP-AD items (% ranging from 55.1 to 79.3) with an overall mean of −.51 (SD = .35) for the location parameters and an overall mean of 1.76 (SD = .31) for the discrimination parameters. The values of the discrimination parameters ranged from 1.30 to 2.22 with items #8 (I have had money problems because of my use), #14 (My use has damaged my social life, popularity, or reputation), and #13 (My use has gotten in the way of my growth as a person) having the greatest overall discriminatory ability across the latent negative consequences severity continuum. In addition, the values of the location parameters ranged from −.15 to −.99 with items #14 (My use has damaged my social life, popularity, or reputation), #9 (My physical appearance has been harmed by my use), and #11 (A friendship or close relationship has been damaged by my use) providing the highest values.

3.6 Total Information Curve

The total information curve provides information on the reliability of the SIP-AD across the range of latent-trait scores and is computed via a combination of the item location and discrimination parameters. As noted by Neal and colleagues (2006), when the total information curve is generally peaked it indicates the highest degree of reliability the scale has at that level of the underlying latent trait score. As shown in Figure 1, the total information curve indicates that the refined SIP-AD is more reliable with scores towards the lower end of the latent negative consequences severity continuum.

3.7 Additional Reliability and Validation Analyses of the SIP-AD

Additional analyses focused on examining the reliability and validity of the refined SIP-AD (i.e., 10 items) via classical measurement techniques. For these analyses, CTT statistics were conducted for the original 15-item SID-AD and the refined 10-item SIP-AD based on the original scoring scheme (i.e., 0 = none to 3 = daily) in order to provide justification for a shortened SIP-AD. The refined 10-item SIP-AD demonstrated a high degree of internal consistency reliability (Cronbach's Alpha = .956). As shown in Table 3, Cronbach's alpha with each item removed were conducted for each of the items, which all were in the high range with relatively little variation (Cronbach's alphas ranging from .949 to .961). Along these lines, item to total scale correlations were conducted with correlations ranging from .696 to .862. With respect to the original 15-item SIP-AD, Cronbach's alpha with each item removed and item to scale correlations revealed negligible differences in comparison to the reliability estimates obtained from the refined 10-item SIP-AD (See Table 3 for comparisons between the 15-item and refined 10-item SIP-AD).

In order to demonstrate the validity of the original 15-item SIP-AD and the refined 10-item SIP-AD, correlation analyses between the each of the SIP-AD measures and other meaningful variables were conducted. Results indicated that original 15-item SIP-AD scores were significantly related to any past substance abuse treatment ($r = .37$, $p < .001$), any self-help attendance ($r = .35$, $p < .001$), frequency of total club drug use ($r = .29$, $p < .001$), cocaine use ($r = .26$, $p < .001$), and crystal methamphetamine use ($r = .20$, $p < .001$). The correlation between number of DSM-IV symptoms endorsed and the original 15-item SIP-AD was high ($r = .54$, $p < .001$). With respect to alcohol use, the original 15-item SIP-AD was not meaningfully related to frequency of alcohol use ($r = .09$, $p > .05$), but it was significantly related to average standard drinks per week ($r = .22$, $p < .001$).

Along these lines, results indicated that refined 10-item SIP-AD sum scores were similar to those found with the 15-item SIP-AD and were significantly related to, albeit with fairly low correlations, any past substance abuse treatment ($r = .37$, $p < .001$), any self-help attendance ($r = .36$, $p < .001$), frequency of total club drug use ($r = .28$, $p < .001$), cocaine use ($r = .24$, $p < .001$), and crystal methamphetamine use ($r = .21$, $p < .001$). The correlation between number of DSM-IV symptoms endorsed and the refined 10-item SIP-AD ($r = .54$, $p < .001$) was

somewhat higher, and was strongly associated with the original 15-item SIP-AD (r = .99, p < . 001). With respect to alcohol use, the refined 10-item SIP-AD was not meaningfully related to frequency of alcohol use (r = .08, p > .05), but it was significantly related to average standard drinks per week (r = .21, p < .001). Lastly, an exploratory components analyses was conducted on the refined 10-item SIP-AD, with a single dominant factor emerging (i.e., factor loadings ranged from .734 to .875) accounting for 67.54% of the variance, providing evidence of its unidimensional structure.

4. Discussion

Current psychometric evaluation of the SIP-AD has been limited to classical test theory (CTT) techniques. Methods from IRT, however, offer several advantages in comparison to methods from CTT such as providing information on the performance of individual scale items as well as determining an instrument's reliability across the underlying trait that it is purported to measure. Recently, these methods have been used to refine measures of alcohol-related problems (Kahler et al., 2003; Neal et al., 2006). The application of IRT to measures that assess simultaneously both alcohol and illicit drug use, such as the SIP-AD, has been limited, and represents a novel direction in addictions research and measurement (Kahler et al., 2003).

Based on the IRT DIF analyses, three items on the SIP-AD were found to be biased with respect to certain sub-groups among MSM. Explanations for why these specific items demonstrated such difference remain unclear. Some suggestions as to why two of these items may be biased to these attributes are explored here. One of the items that demonstrated significant DIF on the location parameters for race/ethnicity ("I have lost weight/not eaten properly") may reflect cultural differences among MSM in perspectives on eating, dietary regimens, and body image. Studies of primarily White individuals have demonstrated that homosexual men score significantly higher pathological scores on scales measuring eating attitudes, symptoms of bulimia, and body image than heterosexual men (Russell & Keel, 2002). While studies investigating the prevalence of eating disorders among gay and bisexual men across ethnicity and race have not demonstrated significant differences (Feldman & Meyer, 2007), some evidence suggests that White gay and bisexual men may use drugs more often than non-whites to control for weight or physical appearance (Halkitis, Fischgrund, & Parsons, 2005). This may indicate that White, or in this case, non-Black, MSM may not identify weight loss or lack of eating properly as a negative consequence of drug use. Instead, they may see drug use as a tool to lose weight.

The item that demonstrated significant DIF on the location parameters for HIV status ("my family has been hurt by my drug use") may be biased due to the extreme psychosocial and medical ramifications of the virus itself. The now extended trajectory of the course of HIV and AIDS can indicate a lengthened need for informal caregiving by family (Wight, Aneshensel, & LeBlanc, 2003). Because it is well established that recreational drug use can severely negatively impact prescription medications used to treat HIV (Grov, Bimbi, Nanin, & Parsons, 2006) and because many family members of HIV positive men may also be men living with HIV (Wight et al., 2003), continued drug use by an individual may be perceived by his family or himself as self-destructive, rather than solely recreational. As a result, HIV positive MSM may feel a heightened sense that their drug or alcohol use negatively impacts their family. Regardless of how plausible these explanations for the racial differences and the differences among HIV positive versus HIV negative MSM may be, further research is necessary to examine potential factors that lead to this bias.

Results from the final IRT model indicated that that the refined 10-item SIP-AD provides greater information at the lower end of the alcohol and illicit drug use negative consequences latent-trait continuum. In other words, the item location and discrimination parameters have a

greater degree of reliability at the lower end of the continuum and the location parameters do not demonstrate a progression from less to more severe alcohol and illicit drug related consequences. This may be explained, in part, by the fact that each of the items may be providing redundant information (i.e., each item has a high probability of endorsement) or that non-treatment seeking MSM represent a high-risk population, who engage in multiple risk factors that lead them to be more likely to experience a greater number of negative consequences. **Continued research is warranted to replicate the current findings across a number of different populations in order to determine the stability of the IRT parameters derived from this sample**.

An additional aim of this study was to examine the reliability and validity of the SIP-AD among a non-treatment seeking MSM sample using traditional CTT statistics. Overall, psychometric analyses on the refined 10-item SIP-AD exhibited a high degree of reliability and validity. Internal consistency reliability of the refined 10-item SIP-AD was high, and was comparable to that of the original 15-item SIP-AD. The correlation between the refined 10-item SIP-AD and original 15-item SIP-AD was extremely high, and validity analyses (i.e., correlations) between each of the SIP measures and other relevant variables were similar suggesting that the shortened measure captures the same information as the original scale. Moreover, the refined 10-item SIP-AD demonstrated good concurrent and discriminant validity as well as measures a unidimensional construct, providing evidence that it accurately reflects the construct of negative consequences of alcohol and illicit drug use. These findings indicate that the refined 10-item SIP-AD is both reliable and valid among non-treatment seeking MSM and is consistent with prior findings that have found the SIP-AD to demonstrate a high degree of reliability and validity among both clinical and non-clinical samples (Blanchard et al., 2003; Gillespie et al., 2007). **It remains necessary for future research to examine the psychometric properties of the refined SIP-AD across other alcohol and drug using populations to ensure its utility as a brief standardized clinical measure**.

There were some limitations associated with the current study. First, all IRT analyses were conducted with each item dichotomized, which does not fully reflect the original scoring scheme of the SIP-AD. A more sophisticated polytomous IRT model needs to be employed in order to examine the SIP-AD's multiple-category item response format. Low endorsement ratings for multiple response options, however, can be problematic for estimation of polytomous IRT parameters, which we found among the current sample, providing justification for conducting binary IRT analyses. Second, the majority of the measures were based on self-report. Assurances of anonymity and multiple indicators of key constructs, all of which were part of this study, can maximize the accuracy of self-reported alcohol and illicit drug use data (Del Boca & Darkes, 2003). Lastly, it is important to note that the current study findings are considered exploratory and continued replication is necessary in order to determine the stability of current study findings.

The addictions field currently lacks standardized instruments that assess recent negative consequences, particularly measures that assess consequences associated with the effects of both alcohol and illicit drug use. In response, the 15-item SIP-AD was developed as a brief measure to be used in both clinical and research settings. The present study findings indicate that the original 15-item SIP-AD can be shortened to 10-items while maintaining a high degree of reliability and validity. As a brief measure, the refined 10-item SIP-AD offers several advantages. Its brevity can reduce time, effort and expenses in administration. In addition, it can be incorporated into a number of health care environments (e.g., physician setting) with relative ease. With respect to its strong psychometric properties, its utility as a uniform measure of alcohol and illicit drug use consequences in research studies will permit comparisons across studies and populations. Within brief interventions, it may serve as a useful tool for providing clients with brief objective feedback regarding their consequences. In summary, if the present

findings are replicated within other studies and populations, the refined 10-item SIP-AD may serve as a non-biased measure for assessing recent negative consequences associated with the use of alcohol and illicit drug use.
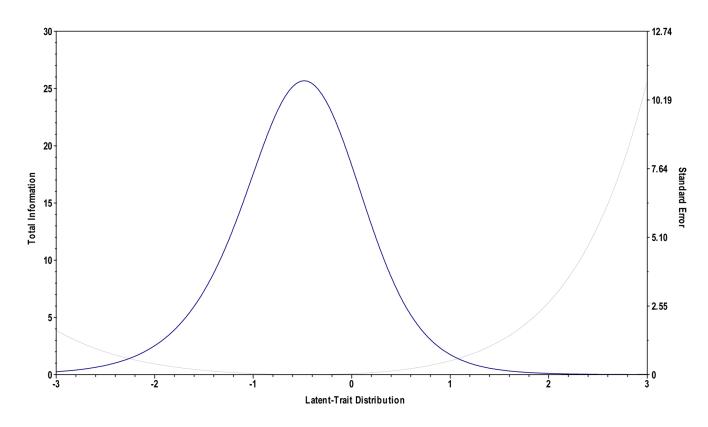
Acknowledgments

References

Bergmark KH. Drinking in the swedish gay and lesbian population. Drug and Alcohol Dependence 1999;56:133–143. [PubMed: 10482404]

Blanchard KA, Morgenstern J, Labouvie E, Morgan T, Bux DA. Assessing consequences of substance use: Psychometric properties of the inventory of drug use consequences (InDUC-2R). Psychology of Addictive Behaviors 2003;17:328–331. [PubMed: 14640829]

Bux DA. The epidemiology of problem drinking in gay men and lesbians: A critical review. Clinical Psychology Review 1996;16(4):277–298.

Cochran SD, Ackerman D, Mays VM, Ross MW. Prevalence of non-medical drug use and dependence among homosexually active men and women in the US population. Addiction 2004;99:989–998. [PubMed: 15265096]

Del Boca FK, Darkes J. The validity of self-reports of alcohol consumption: State of the sceince and challenges for research. Addiction 2003;98(2):1–12. [PubMed: 14984237]

Embretson, SE.; Reise, SP. Item Response Theory for Psychologists. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.; 2000.

Feldman MB, Meyer IH. Eating disorders in diverse lesbian, gay, and bisexual populations. International Journal of Eating Disorders 2007;40:218–226. [PubMed: 17262818]

Ferrando S, Goggin K, Sewell M, Evans S, Fishman B, Rabkin J. Substance use disorders in gay/bisexual men with HIV and AIDS. The American Journal on Addictions 1998;7:51–60. [PubMed: 9522007]

Gillespie W, Holt JL, Blackwell RL. Measuring outcomes of alcohol, marijuana, and cocaine use among college students: A preliminary test of the Shortened Inventory of Problems--Alcohol and Drugs (SIP-AD). Journal of Drug Issues 2007;37:549–568.

Grov C, Bimbi DS, Nanin JE, Parsons JT. Exploring racial and ethnic differences in recreational drug use among gay and bisexual men in New York City. Journal of Drug Education 2006;36(2):105–123. [PubMed: 17153512]

Grov C, Bux DA, Parsons JT, Morgenstern J. Recruiting hard-to-reach men who have sex with men into an intervention study: Lessons learned and implications for applied research. Substance Use and Misuse. (in press).

Halkitis PN, Fischgrund BN, Parsons JT. Explanations for methamphetamine use among gay and bisexual men in New York City. Substance Use and Misuse 2005;40:1331–1345. [PubMed: 16048820]

Harawa NT, Williams JK, Ramamurthi HC, Manago C, Avina S, Jones M. Sexual behavior, sexual identity, and substance abuse among low income bisexual and non-gay-identifying African American men who have sex with men. Archives of Sexual Behavior 2008;37:748–762. [PubMed: 18546069]

Kahler CW, Strong DR, Hayaki J, Ramsey SE, Brown RA. An item response analysis of the Alcohol Dependence Scale in treatment-seeking alcoholics. Journal of Studies on Alcohol 2003;64:127–136. [PubMed: 12608493]

Klitzman RL, Greenberg JD, Pollack LM, Dolezal C. MDMA ('ecstasy') use and its association with high risk behaviors, mental health, and other factors among gay/bisexual men in New York City. Drug and Alcohol Dependence 2002;66:115–125. [PubMed: 11906799]

Klitzman RL, Pope HG, Hudson JI. MDMA ('ecstasy') abuse and high risk sexual behaviors among 169 gay and bisexual men. American Journal of Psychiatry 2000;157:1162–1164. [PubMed: 10873928]

Koblin BA, Husnik MJ, Colfax G, Huang Y, Madison M, Mayer K, et al. Risk factors for HIV infection among men who have sex with men. AIDS 2006;20:731–739. [PubMed: 16514304]

Miller, WR.; Tonigan, JS.; Longabaugh, R. The Drinker Inventory of Consequences (DrInC): An instrument for assessing adverse consequences of alcohol abuse. Test manual. Rockville, MD: NIAAA Project MATCH Monograph Series Volume 4; 1995.

Neal DJ, Corbin WR, Fromme K. Measurement of alcohol-related consequences among high school and college students: Application of item response models to the rutgers alcohol problem index. Psychological Assesment 2006;18(4):402–414.

Nollen NL, Goggin K, Heiland MF, Catley D, Kessler K, Brammer SK, et al. The impact of combination therapy on health behaviors of HIV+ gay men. AIDS and Behavior 2002;6(2):131–139.

Parsons JT, Halkitis PN, Bimbi DS. Club drug use among young adults frequenting dance clubs and other social venues in New York City. Journal of Child & Adolescent Substance Abuse 2006;15(3):1–14.

Russell CJ, Keel PK. Homosexuality as a specific risk factor for eating disorders in men. International Journal of Eating Disorders 2002;31:300–306. [PubMed: 11920991]

Scientific Software International, I. Parscale 4.1. Lincolnwood, IL: Author; 2003.

Stall R, Paul JP, Greenwood G, Pollack LM, Bein E, Crosby GM, et al. Alcohol use, drug use and alcohol-related problems among men who have sex with men: The Urban Men's Health Study. Addiction 2001;96(11):1589–1601. [PubMed: 11784456]

Tonigan JS, Miller WR. The Inventory of Drug Use Consequences (InDUC): Test-retest stability and sensitivity to detect change. Psycholology of Addictive Behaviors 2002;16(2):165–168.

Torian LV, Makki HA, Menzies IB, Murrill CS, Weisfuse IB. HIV infection in men who have sex with men, New York City Department of Health sexually transmitted disease clinics, 1990–1999: A decade of serosurveillance finds that racial disparities and associations between HIV and gonorrhea persist. Sexually Transmitted Diseases 2002;9:73–78. [PubMed: 11818891]

Tourangeau R, Smith TW. Asking sensitive questions: The impact of data collection mode, question format, and question context. Public Opinion Quarterly 1996;60:275–304.

Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH. Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. Science 1998;280:867–873. [PubMed: 9572724]

Vaudrey J, Raymond HF, Chen S, Hecht J, Ahrens K, McFarland W. Indicators of use of methamphetamine and other substances among men who have sex with men, San Francisco 2003–2006. Drug and Alcohol Dependence 2007;90(1):97–100. [PubMed: 17428622]

Wheeler DP, Lauby JL, Liu K, Van Sluytman LG, Murrill CS. A comparative analysis of sexual risk characteristics of black men who have sex with men or with men and women. Archives of Sexual Behavior 2008;37:697–707. [PubMed: 18509753]

Wight RG, Aneshensel CS, LeBlanc AJ. Stress buffering effects of family support in AIDS caregiving. AIDS Care 2003;15(5):595–613. [PubMed: 12959815]

Wright, EM. Substance abuse in African American communities. In: Straussner, SLA., editor. Ethnocultural factors in substance abuse treatment. New York: The Guilford Press; 2001. p. 31-51.

Figure 1.
Total information curve for the SIP-AD. The vertical axis to the left is the total test information for a given scale score and is represented by the dark curvilinear line. The vertical axis to the right is the standard error for a given scale score and the measurement error of the scale is plotted by the dotted line. The horizontal line represents the level of ability that is measured by the underlying construct (i.e., negative consequences severity). This figure was generated from PARSCALE 4.1 software.

Table 1

Differential item functioning analyses according to race/ethnicity (AA vs. Non-AA) and HIV status (HIV negative vs. HIV positive) for the location and slope parameters

Item	% Endorsement		Location Parameters		Chi-square	% Endorsement		Location Parameters		Chi-square
	AA	Non-AA	AA	Non-AA	$X^2(1)$	HIV positive	HIV negative	HIV positive	HIV negative	$X^2(1)$
10) My family has been hurt by my use	62.9	41.8	−.338	−.181	1.125	58.5	35.7	−.249	.097	5.191[*]
14) My use has damaged my social life, popularity, or reputation	64.1	49.7	−.413	−.417	.001	62.4	43.4	−.371	−.148	2.552
9) My physical appearance has been harmed by my use	66.3	54.1	−.507	−.556	.104	64.1	50.1	−.451	−.321	.779
4) When using my personality has changed for the worse	69.1	60.9	−.583	−.766	1.319	64.5	63.2	−.461	−.715	2.936
6) While using, I have said harsh or cruel things to someone.	71.4	55.8	−1.044	−.624	2.905	63.4	58.8	−.522	−.644	.198
11) A friendship or close relationship has been damaged by my use	72.1	52.4	−.682	−.511	1.048	64.1	52.7	−.444	−.405	.068
1) I have been unhappy because of my use	73.1	71.8	−.894	−1.228	3.193	75.6	67.1	−.961	−.881	.168

	% Endorsement		Location Parameters		Chi-square	% Endorsement		Location Parameters		Chi-square
	AA	Non-AA	AA	Non-AA	$X^2(1)$	HIV positive	HIV negative	HIV positive	HIV negative	$X^2(1)$
12) I have lost interest in activities and hobbies because of my use	75.4	53.7	-.849	-.552	3.592[*]	65.9	55.5	-.498	-.484	.011
3) I have failed to do what is expected of me because of my use	77.1	63.3	-.865	-.831	.041	71.1	64.3	-.711	-.738	.037
2) Because of my use, I have lost weight or not eaten properly	78.3	54.3	-1.091	-1.471	3.511[*]	78.4	78.1	-1.09	-1.342	1.494
13) My use has gotten in the way of my growth as a person	78.3	58.5	-.911	-.703	1.636	70.7	58.2	-.666	-.561	.531
8) I have had money problems because of my use	80.1	69.1	-1.016	-1.003	.007	78.1	65.4	-.934	-.761	1.281
15) I have spent too much or lost a lot of money because of my use	82.3	71.4	-1.104	-1.098	.001	77.4	72.5	-.929	-.998	.191
7) When using, I have done impulsive things that I regretted later	82.9	77.2	-1.16	-1.301	.486	79.1	79.7	-1.041	-1.288	1.854
5) I have taken foolish	83.4	81.6	-1.354	-1.572	1.02	82.9	81.1	-1.273	-1.441	.625

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

% Endorsement		Location Parameters		Chi-square	% Endorsement		Location Parameters		Chi-square
AA	Non-AA	AA	Non-AA	$X^2(1)$	HIV positive	HIV negative	HIV positive	HIV negative	$X^2(1)$
risks when I have been using									

Note: The current table is sort in descending sequence according to AA's location parameters

*
denotes p < .05

**
denotes p < .01

Table 2

Final IRT model for the refined SIP-AD based on the remaining 10-items

	% Endorsement of item (N)	Location	S.E.	Discrimination	S.E.
14) My use has damaged my social life, popularity, or reputation	55.0 (258)	−.15	.07	2.06	.25
9) My physical appearance has been harmed by my use	58.6 (275)	−.26	.07	1.70	.20
11) A friendship or close relationship has been damaged by my use	59.7 (280)	−.29	.07	1.72	.21
4) When using my personality has changed for the worse	64.0 (300)	−.44	.07	1.54	.19
13) My use has gotten in the way of my growth as a person	65.9 (309)	−.47	.07	2.05	.27
3) I have failed to do what is expected of me because of my use	68.4 (321)	−.58	.08	1.60	.19
8) I have had money problems because of my use.	73.1 (343)	−.69	.07	2.22	.33
1) I have been unhappy because of my use	72.3 (339)	−.75	.09	1.30	.15
15) I have spent too much or lost a lot of money because of my use	75.5 (354)	−.78	.08	1.95	.26
7) When using, I have done impulsive things that I regretted later	79.3 (372)	−.99	.10	1.50	.21
		Mean = −.51 (SD = .35)		Mean = 1.76 (SD =.31)	

Note: The items in the chart are sorted in descending sequence according to the location parameters

Table 3

Item to Total Scale Correlations and Cronbach's Coefficient Alpha with each item removed for each of the original 15-item SIP-AD and refined 10-item SIP-AD

	Item-Total Scale Correlation	Cronbach's Alpha if Item Deleted	Item-Total Scale Correlation	Cronbach's Alpha if Item Deleted
	Original 15-item SIP-AD	Original 15-item SIP-AD	Refined 10-item SIP-AD	Refined 10-item SIP-AD
1) I have been unhappy because of my use	.691	.967	.696	.961
2) Because of my use, I have lost weight or not eaten properly	.769	.965	*	*
3) I have failed to do what is expected of me because of my use	.828	.964	.820	.951
4) When using my personality has changed for the worse	.797	.965	.776	.952
5) I have taken foolish risks when I have been using	.804	.965	*	*
6) While using, I have said harsh or cruel things to someone	.674	.967	*	*
7) When using, I have done impulsive things that I regretted later	.783	.965	.750	.953
8) I have had money problems because of my use	.849	.964	.855	.951
9) My physical appearance has been harmed by my use	.816	.964	.819	.950
10) My family has been hurt by my use	.753	.966	*	*
11) A friendship or close relationship has been damaged by my use	.811	.964	.792	.952
12) I have lost interest in activities and hobbies because of my use	.851	.964	*	*
13) My use has gotten in the way of my growth as a person	.865	.963	.862	.949
14) My use has damaged my social life, popularity, or reputation	.843	.964	.839	.950
15) I have spent too much or lost a lot of money because of my use	.853	.964	.857	.949

	Item-Total Scale Correlation	Cronbach's Alpha if Item Deleted	Item-Total Scale Correlation	Cronbach's Alpha if Item Deleted
	Original 15-item SIP-AD	Original 15-item SIP-AD	Refined 10-item SIP-AD	Refined 10-item SIP-AD
Overall SIP-AD		.967		.956

Note: * denotes that the item was dropped for the present reliability analyses