



Published in final edited form as:

J Drug Issues. 2010 March ; 40(2): 495–516.

Self-report of Longitudinal Substance Use: A Comparison of the UCLA Natural History Interview and the Addiction Severity Index

Debra A. Murphy, Ph.D.¹, Yih-Ing Hser, Ph.D.², David Huang, Dr.Ph.H.³, Mary-Lynn Brecht, Ph.D.⁴, and Diane M. Herbeck, M.A.⁵

¹ Health Risk Reduction Projects, Integrated Substance Abuse Programs, Department of Psychiatry University of California at Los Angeles, 11075 Santa Monica Blvd., Suite 200, Los Angeles, CA 90025 dmurphy@mednet.ucla.edu, (310) 794-8127

² Integrated Substance Abuse Programs, Department of Psychiatry, University of California at Los Angeles, 1640 S. Sepulveda Blvd, Suite 200, Los Angeles, CA 90025, yhser@ucla.edu, (310) 267-5388

³ Integrated Substance Abuse Programs, Department of Psychiatry, University of California at Los Angeles, 1640 S. Sepulveda Blvd, Suite 200, Los Angeles, CA 90025, yhuang@ucla.edu, (310) 267-5288

⁴ Integrated Substance Abuse Programs, Department of Psychiatry, University of California at Los Angeles, 1640 S. Sepulveda Blvd, Suite 200, Los Angeles, CA 90025, lbrecht@ucla.edu, (310) 267-5275

⁵ Integrated Substance Abuse Programs, Department of Psychiatry, University of California at Los Angeles, 1640 S. Sepulveda Blvd, Suite 200, Los Angeles, CA 90025, dherbeck@ucla.edu, (310) 267-5329

Abstract

Reliance on self-report of alcohol and drug use behavior is typical among studies of substance abusers. Few studies have compared different instruments assessing frequency of drug use over long periods of time to compare findings and determine if the pattern of use is shown to be similar across measures. In this study, the UCLA Natural History Interview (NHI) and the Addiction Severity Index (ASI) were administered at three annual follow-up periods ($N = 301$). The temporal pattern of the trajectories of days of use assessed by the ASI and NHI are comparable (in terms of both slope and intercept) for alcohol, heroin, cocaine, methamphetamine, and marijuana use. Some discrepancies appear to arise from differences in terminology among the instruments. However, the patterns of drug use were consistent across instruments, supporting their reliability for longitudinal examination of self-reported drug use.

Keywords

Self-report; Substance Use; Validation Studies; Reliability; Natural History Interview; Addiction Severity Index

Introduction

Reliance on self-report of alcohol and drug use behavior by substance users is typical among studies of illicit drug use. Self-report is often the only feasible methodology that can be utilized to address the research questions of interest (Darke, 1998). Numerous substance use surveys challenge participants' cognitive abilities by asking them to recall the date, frequency, or characteristics of personal events, which is a task of episodic memory retrieval (Tulving, 1983). Those phenomena of interest to health-behavioral researchers are often unlikely to be documented in any form (Croyle & Loftus, 1992), particularly a respondent's daily alcohol and drug use throughout their lifetime, and thus self-report may be the only available assessment option. Data on the frequency that individuals use drugs are often utilized to describe the severity of their drug use problems; in clinical settings drug use frequencies before and after treatment are used to establish the effectiveness of an intervention; and drug use frequency information collected in population surveys is used to gauge the prevalence of severe drug use problems (Morral, McCaffrey, & Iguchi, 2000). There are extremely few studies that have compared longitudinal self-report of substance use using multiple instruments. Typically, administering the same instrument two times and including overlapping timeframes is done, but questions remain regarding biases specific to that instrument and the validity of self-report over long timeframes. The current study compared longitudinal self-report of substance use using multiple instruments.

While prospective studies can include concurrent biomarker measures such as urinalysis, costs and subject burden associated with data collection and analysis of such markers over time may be limited and may constrain the possibility of analyses to identify many complexities of longitudinal patterns. In addition, retrospective longitudinal studies of past drug use do not have these measures available. In such retrospective studies, concerns about accuracy of reports of substance use continue to be raised, given the numerous reasons why self-report of drug users might be inaccurate: recollection of events are characterized by particular times, locations, personal relevance, and in many cases by emotion (Croyle & Loftus, 1992); memory for any behavior is likely not to be perfect, especially if assessed over long time periods (Darke, 1998); and response biases in self-report can be attributed to recall methods, recall periods, and operational definitions of questions (Killeen, Brady, Gold, Tyson, & Simpson, 2004). Recall errors occur when a respondent simply cannot remember the correct information requested, and therefore cannot report it accurately; the likelihood of such errors depends on the specificity and time scale of the questions (Kuha, 2001). With respect to retrospective recalls, accuracy of survey reports has been shown to decline with the length of recall required. In examining a recall of lifetime disorders and age of onset of disorders, Prusoff, Merikangas, & Weissman (1988) found decreasing reliability between baseline, 6-month, and 18-month interviews, which calls into question much longer-term follow-up studies. Nevertheless, more frequent assessments require greater resources, which may not always be feasible.

Because reliability is a precondition for validity (Nunnally, 1978), examining the stability and consistency of self-report of substance use is an important step in establishing the value of such assessments. Substance use is one of the cases in which it is clear that errors of measurement are potentially large enough to be of serious concern (Kuha, 2001). For example, in longitudinal studies there is documentation of inconsistent reports, such as a person reporting some lifetime use of a substance at some age but no lifetime use at a later age (Adair, Craddock, Miller, & Turner, 1996; Siddiqui, Mott, Anderson, & Flay, 1999). More emphasis has been placed on trying to establish reliability and consistency of self-reported age of substance use initiation (Johnson & Mott, 2001), rather than on the frequency of use over long time spans. In addition, the majority of studies have tried to utilize a test-retest approach, in which the same instrument is used in an over-lapping

timeframe to determine if similar self-reports are provided. However, few studies have utilized different instruments that assess frequency of drug use over long periods of time to determine whether the pattern of use is shown to be similar across the two measures, which is the focus of the present article.

Literature Review

There are a large number of studies that focus on recent retest of self-reported drug use that utilize biomarkers that were collected at the time of self-report to validate the self-report, such as urinalysis (Anglin, Hser, & Chou, 1993; Cook, Bernstein, Arrington, Andrews, & Marshall, 1995; Murphy, Durako, Muenz, & Wilson, 2000) or hair analysis (Cooper et al., 2000; Fendrich, Johnson, Sudman, Wislar, & Spiehler, 1999; Magura, Freeman, Siddiqi, & Lipton, 1992; Nyamathi, Leake, Longshore, & Gelberg, 2001). These studies indicate somewhat high congruence between self-report and urinalysis/hair analysis, as well as some underreporting. However, these findings will not be covered in detail here, as the focus in this paper is on longitudinal retrospective self-report of frequency of drug use studies for which often these biomarker measures are not available.

A number of studies have evaluated re-interview effects in substance use surveys over long time periods (Fendrich & Vaughn, 1994). As noted earlier, many of these have focused on either age of first use, or whether a participant reports ever vs. never using a drug. Mensch and Kandel (1988) compared responses to questions about marijuana use over two waves that were four years apart, and found that 9% of those initially reporting use indicated that they never used marijuana at follow-up. In an 11-year longitudinal study for self-reported age of onset for alcohol, tobacco, and illicit drug involvement, Parra, O'Neill, & Sher (2003) found a moderate degree of reliability for the three substances.

The level of detail collected about frequency of use varies across studies. For example, in a study of self-report stability for substance use over 10 years, Shillington, Cottler, Mager, & Compton (1995) found that cannabis, cocaine, sedatives, and opiates had high agreement rates (ranging from 82 – 86%). However, participants were simply asked to review a list of substances and asked if they had ever used any of the substances more than five times; agreement was calculated as no use at either time or use at both times. In a study of reliability of recalled alcohol use from 10 years in the past, participants first reported their drinking habits in 1971 – 1975 (Liu et al., 1996). During a follow-up interview in 1982 – 1984, they were asked to recall their drinking usage from 10 years earlier and to also report current drinking levels. The correlation for recalled alcohol intake vs. reported intake at baseline was good ($r = .7$), with current heavier drinkers tending to underestimate their previous amount of drinking. Sobell, Sobell, Leo, & Cancilla (1988) also found that test-retest reliability of alcohol abusers' reports of their past (approximately 8 years prior to the interview) drinking behavior and life events that occurred are generally reliable. Sobell, Kwan, & Sobell (1995) have reported similar findings for drug abusers in treatment. Participants were asked about pre-treatment lifetime drug use and demographic and drug-related events, and good test-retest reliability was found for most reports of drug use and related events.

The consistency of retrospective recall by heroin users using the Natural History Instrument (described in greater detail in the Method section) was investigated in a series of studies (Anglin et al., 1993; Chou, Hser, & Anglin, 1996; Hser, Anglin, & Chou, 1992). Ten-year retest measures of estimated drug use and other behaviors (e.g., crime, drug trafficking, employment) over an overlapping period of 4 years were obtained. Compared to the initial interview, the 10-year follow-up yielded higher levels of retrospective heroin use for the overlapping period than did the initial interview. The correlation for narcotic abstinence over the period was 0.71, and 0.63 for daily narcotic use (Hser et al., 1992). Test-retest

correlations were higher among higher-frequency activities (Anglin et al., 1993). Correlation coefficients of inter-variable relationships, based on 46 variables measured at two interviews 10 years apart, ranged as high as 0.86 and 0.90 (Chou et al., 1996). Consistent with Darke (1998), overall these data indicate that injection drug users (IDUs) give reasonably reliable answers to questions about drug use.

In a study similar to the current investigation, a comparison of the Addiction Severity Index and a Retrospective Life History Calendar was conducted (Lewandowski & Hill, 2003). Looking at percent agreement between the two instruments, the calendar method was more likely to identify more substances and longer durations of use, although there were exceptions within drug categories. Furthermore, the timeline method of gathering retrospective reports of recent drinking were compared with data gathered from the same participants using a common quantity-frequency (QF) method, and the timeline method was shown to have fairly good reliability, whereas the QF method provided a relatively insensitive measure of individual differences in drinking behavior (Sobell et al., 1988). Sobell et al. (1988) report the QF method did not accurately reflect individual drinking patterns, as it categorized almost half of the sample as “high” consumers of alcohol, with no further differentiation within that group.

Summary and Purpose of Present Study

Overall, there exists a large body of literature on reliability of self-reported drug use other than those reviewed thus far (Amsel, Mandell, Mathias, & Hoehnerman, 1976; Darke, Hall, Wodak, & Heather, 1992; Ehrman & Robbins, 1994; Kokkevi, Richardson, Palermou, & Leventakou, 1997). However, it is difficult to find studies that have compared longitudinal self-report of substance use using multiple instruments. Utilizing the same instrument twice for overlapping timeframes seems to be the common method of choice, but the questions remain as to whether biases specific to that instrument result in valid self-report over long timeframes. In the present study, two different instruments, which both assess frequency of self-reported drug use, were administered longitudinally. The UCLA National History Interview (NHI) and the Addiction Severity Index (ASI; both described in detail in the Method section) were administered during three annual follow-up interviews, permitting comparisons of drug use data between the two instruments across multiple time points.

As mentioned earlier, longitudinal data are needed to assess changes in drug use for various purposes. These longitudinal data can be collected either prospectively via repeated measures of current use, taken at multiple time points, or retrospectively via recall of past use over a long period of time. Self-report of current use has the advantage of fewer memory problems due to recency of the event, but may be subject to biased reporting, particularly if there are perceived immediate consequences associated with drug use (e.g., over-reporting at treatment admission for desired services, under-reporting among criminal justice participants to avoid sanction). Retrospective recall is more likely to be influenced by memory problems, but is more likely to solicit honest reporting and is less costly. Scientifically it is important to learn if findings converge or diverge using the two approaches and different instruments while aiming to measure the same behaviors during the specified time points or periods. The purpose of the present study is to compare drug use levels collected by two instruments administered multiple times and requiring either prospective reporting of current use or retrospective recall of past use over multiple years.

Method

Study Design and Sample

The study sample was part of the Treatment Utilization & Effectiveness Study (TUE), which is a prospective longitudinal study designed to assess the nature and extent of drug treatment utilization and effectiveness among a range of high-risk populations. In 1992 – 1994, more than 5,000 men and women were interviewed and screened in sexually transmitted disease (STD) clinics, emergency rooms (ER), and jails in Los Angeles County (Hser, Boyle, & Anglin, 1998). Follow-up interviews 1, 2 and 3 were conducted at 12, 24 and 36 months, respectively, after the baseline interview, with successively smaller random subsets of the original sample (i.e., approximately 1,800 drug users were identified at study intake, 930 drug users were selected and interviewed at Follow-up 1, 566 at Follow-up 2, and 304 at Follow-up 3). Since the dynamic forms section of the Natural History Interview (NHI) was only conducted at follow-up 2 and 3, only 566 participants who had completed both the static and dynamic forms sections of the NHI and the ASI (these measures are described below) were considered for inclusion in the final research sample. Of these 566 participants, 301 participants completed all three follow-up interviews and are included in the analyses for this paper. The sample consisted primarily of African Americans (53.8%), Hispanics (25.9%), and Whites (17.6%), and 61.5% were male. At study intake, the age of the participants averaged 31.6 years, and the majority of the participants were single (59.8%). Approximately 20% of the participants had less than high school education and 71.8% were not employed. See Table 1 for participant demographics. Attrition analysis was conducted, comparing the study sample of 301 participants with the 265 participants who were not included in this analysis. Results show no demographic differences (i.e., gender, race/ethnicity, marital status, education, employment status), or differences in rates of drug use between the study sample and those excluded. One difference between the two groups was observed: the study sample included a slightly larger proportion of young adults age 18–24 (28.2% compared to 21.3%, $p = 0.04$). However, a comparison of mean ages indicates there was not a significant age difference between the two groups (study sample mean age = 31.6, $SD = 8.6$ vs. excluded participants mean age = 32.9, $SD = 9.1$). These results indicate demographics and drug use were similar between the study sample and those excluded, thus, the exclusion of participants from the study sample does not appear to have introduced a bias.

Data Collection Procedures

At baseline, study participants who identified themselves as users of illegal drugs participated in a structured, face-to-face interview conducted in a private setting by trained UCLA interviewers. Interviewers had Bachelor's degrees or several years of experience conducting research interviews in other studies of a similar nature. Participants were informed that the purpose of the study was to assess drug treatment use and effectiveness and to suggest ways to improve treatment. Participants in STD clinics and emergency rooms received token monetary compensation while those in jails received food as compensation, and the interview took approximately 45 minutes to complete. A voluntary urine specimen was collected from non-incarcerated participants at the end of the interview; urine samples were collected from 52.5% ($n = 158$) of participants at follow up 1; 43.9% ($n = 132$) at follow up 2; and 46.2% ($n = 139$) at follow up 3.

Each subsequent face-to-face follow-up interview took approximately 2 – 3 hours and participants were paid for the interview. At the end of the interview, a voluntary urine specimen was collected from those who were not incarcerated. Respondents were shown a copy of the confidentiality certificate obtained from DHHS under PL 94-255 for exemption

of the data from any legal proceedings, and they were assured that all information provided to the researchers would be held in strictest confidence.

Instruments and Measures

Two questionnaires (the UCLA National History Interview and Addiction Severity Index) were administered during the annual follow-up interview, permitting comparisons of drug use data between the two instruments. Although the NHI and ASI include similar drug use measures, the instruments assess drug use in different ways and within different contexts as described below.

UCLA Natural History Interview (NHI)—The NHI has been used for more than the past two decades for the evaluation of civil commitment and methadone maintenance treatment for narcotics addicts (Anglin & McGlothlin, 1984; McGlothlin, Anglin, & Wilson, 1977) and has been used with various drug-abusing populations. The instrument was adapted from that originally designed by Nurco, Bonito, Lerner, & Balter (1975) and was designed to collect retrospective longitudinal quantitative data on drug use and related behaviors. Research in recent years demonstrates that this interview can be easily and effectively adapted and modified to accommodate the entire spectrum of drug use problems and treatment modalities. The detailed natural history interview yields rich information about patterns of drug use and treatment over the course of an addiction career for a period from 12 months before first use to the time of interview.

The instrument consists of a set of “static” and a set of “dynamic” forms that permit the capture of longitudinal, sequential data on drug use, employment, criminal involvement, treatment, and other behaviors over the life course of the participants (McGlothlin et al., 1977). The static forms collect background information on the participant and are administered once during the interview. The dynamic forms are used to collect retrospective and current data on the drug-use history of the participants as well as data on events that might have shaped or have been shaped by drug use (e.g., crime, incarceration, employment, social support network, medical status, psychiatric status, drug treatment). The dynamic part of the interview consists of the repeated administration of these forms for as many life segments (defined by major changes in behaviors or life events being assessed) as necessary. The procedure requires that the interviewer work closely with the respondent to structure the periods of interest, using corroborative information and memory aids (e.g., major life events, historical events). In this way, drug use, criminal behavior, and periods of legal supervision and treatment participation are anchored to major life events, such as the birth of a child, death of a family member, move to a new location, or loss of a job. As described in the literature review, the NHI has demonstrated high congruence with urinalysis and good test-retest reliability.

The NHI static form asked, “How many days have you used [a list of alcohol and drugs] in the last 4 weeks?” The form was administered at each of the three follow-up interviews. The dynamic forms were used in interviews at Follow-up 2 and Follow-up 3. During the Follow-up 2 interview, the dynamic forms collected the monthly record of substance use since age at first use for each of specified types of drugs, including narcotics, cocaine/crack, meth/amphetamine/speed/crystal, and marijuana, and alcohol. The dynamic forms include a separate form for each specified substance. A timeline is created for each of these substances to indicate periods of homogeneous use (e.g., no use, infrequent use [1–4 days per month], frequent use [5 days per month to 5 days per week], and regular use [6 to 7 days per week]). Study participants were asked, “For the period [from month/year to month/year], how often did you use [the substance] during this period?” The response categories were daily (6 days/week), days/week, weekends/month, days/month, days/year, and days/segment. Responses

were converted to the number of days during the month of interest. During the Follow-up 3 interview, the history of monthly drug use was updated to the time of that follow-up.

Addiction Severity Index—The Addiction Severity Index (ASI) is a structured interview that is one of the most widely used substance abuse assessment instruments, particularly in addiction treatment and research (McLellan et al., 1985). It assesses problem severity, both in the past 30 days and lifetime, in seven areas: alcohol use, drug use, employment, family and social relationships, legal, psychological, and medical status. Regarding alcohol and drug use, participants were asked to report “the number of days in the past 30 days that [a list of alcohol and drugs] was used.” Appleby, Dyson, Altman, & Luchins (1997) reported standardized Cronbach alphas of .89 for the alcohol scale, and .79 for the drug use scale. Alterman, Brown, Zaballero, & McKay (1994) also found fairly high internal consistency for the ASI drug and alcohol scales with .62 for the drug area and .87 for the alcohol scale. The ASI was administered at each follow-up.

Drug Use Variables for Analysis

Our analyses focus on the use of alcohol, heroin, cocaine, meth/amphetamine, and marijuana—the five prevalent substances used by substance abusers in treatment or in the general population. The exact name of the specific drug varies somewhat across instruments. The ASI asks for use in the past 30 days of alcohol (any level of use), heroin, cocaine, amphetamine, and cannabis. The NHI-static form asks for use in the last four week of alcohol, heroin, crack, cocaine, amphetamine or any other speed (including methamphetamine), and marijuana or hashish. The NHI-dynamic form asks for alcohol, narcotics, cocaine/crack, amphetamine/speed/crystal, and marijuana use for each month. Note that we will use the label “past month” in this paper to cover all three slightly different time periods. Note also that the NHI-static form had crack and cocaine separately; we thus added them together as cocaine/crack use. Note that for convenience of presentation, we use the simplified label “meth” (or “methamphetamine”) in the remainder of this report. This is consistent with the predominance of use of meth over other amphetamines in many parts of the U. S., as evidenced by the predominance of methamphetamine (23%) among primary drug listed for treatment admissions in Los Angeles County in 2007, as opposed to less than 1% for other amphetamines (Brecht, 2008).

For analyses, indicators measured drug use as number of days with use of the specified drug in the past month. The NHI static form provided three measures, representing use in the four weeks preceding each follow-up interview. Individual items from the ASI also provided indicators at the same three time points for past 30-day use. The NHI dynamic form provided indicators of monthly use, from the longer histories; for these analyses a three-year period was extracted beginning at the first follow-up and continuing through follow-up 3, to correspond to the time periods covered by the NHI static form and ASI. For correlation analyses, three specific months were extracted from the NHI-dynamic histories corresponding to each follow-up.

Congruence between reported drug use on the NHI and ASI with urinalysis

Urine tests were compared to self reported drug use in the past 30 days on the NHI static and dynamic forms and the ASI at each follow-up, indicating a fairly high congruence between self-report and urinalysis for most substances. Congruence for heroin use ranged from 92.4% to 95.4%; for cocaine, congruence ranged from 77.9% to 86.7%; for methamphetamine, congruence ranged from 93.7% to 97.1%; for marijuana, congruence ranged from 70.6% to 79.0%.

Analytic Methods

Pair-wise congruence among the three sources (NHI-static form, NHI-dynamic form, and ASI) of drug use measures (number of days of the specified drug use each month) during the corresponding observation time points were assessed using correlations (for relative consistency; participants who reported no use of a substance were excluded from that specific analysis.) and t-tests (for similarity of level of reported use). The congruence of longitudinal trajectories was assessed using mixed models, estimated using SAS PROC MIXED to test differences in intercepts or slopes across the three sources of measures for drug use trajectories.

Results

Table 2 shows correlation coefficients of days of alcohol/drug use in the past month between each pair of measures across the three instrument sources. Correlations between the ASI and the NHI-static form are consistently high. Except for a correlation of 0.65 on methamphetamine use at follow-up 2, correlations between ASI and NHI-static form indicators at all time points are at or greater than 0.80. Correlations between the NHI-dynamic and the ASI form indicators and between the NHI-dynamic and the NHI-static form indicators range from 0.60 to 0.97 at follow-up 2 and 3. But correlations for measures at follow-up 1 are systematically lower, ranging from 0.14 to 0.52.

The consistency across instruments in the level of self-reported alcohol/drug use was examined at each follow-up using paired t-tests (see Figures 1 – 5 in which means are plotted). No significant differences are seen for mean number of days use in past month for alcohol, heroin, or meth. For cocaine, a significant difference ($p < .01$) is seen at follow-up 3 for NHI-dynamic (mean 2.85) compared to NHI-static (1.25); ASI (1.32) was similar to NHI-static, but the ASI difference from NHI-dynamic was not significant. For marijuana, we see a significant ($p < .01$) difference at follow-up 1 for NHI-dynamic (3.29) compared to ASI (2.38) and to NHI-static (2.47).

The plots in Figures 1 – 5 represent the observed patterns over time for each type of drug based on indicators from the NHI-static form, NHI-dynamic form, and ASI during the corresponding observation time points. Month 0 in each figure corresponds with follow-up interview 1, month 12 with follow-up 2, and month 24 with follow-up 3. Examining these plots visually, we see that trajectories of days of use from the ASI and NHI-static forms are very close to each other for each alcohol/drug. Trajectories of use of the 4 types of drugs (heroin, cocaine, meth, marijuana) measured by the NHI-dynamic form are slightly higher than those from the other sources. For cocaine and meth, these differences (between NHI-dynamic and ASI/NHI-static) become somewhat larger at month 24 (follow-up 3). Trajectories of alcohol use are comparable between NHI-dynamic and ASI/NHI-static forms from month 0 to month 12; however, by month 24 (follow-up 3) the alcohol indicator from the NHI-dynamic form becomes higher than those from the other two sources. Based on estimates from the mixed model for each drug, parameters representing slope and intercept differences across data sources were not significant. For example, for cocaine the parameter estimates were the following: ASI to NHI difference in intercept, -0.71 (standard error [SE] = 0.46, $p = .128$); static to NHI difference in intercept, -0.64 ($SE = 0.46$, $p = .162$); ASI to NHI difference in slopes, -0.01 ($SE = 0.02$, $p = .553$); static to NHI difference in slopes, -0.02 ($SE = 0.02$, $p = .402$). Estimates for other drugs followed a similar pattern. Thus, we conclude that in patterns for each drug, intercepts across data sources were similar and slopes were similar.

Discussion

The temporal pattern of the trajectories of days of use assessed by the ASI and NHI are comparable (in terms of both slope and intercept) for alcohol, heroin, cocaine, methamphetamine, and marijuana use. This indicates that while the number of days of use may not have been reported at exactly the same value, the overall pattern of alcohol and/or drug use across time is being represented in a consistent manner across the assessment instruments. In spite of isolated discrepancies at a specific follow-up for two drugs, the level of self-reported alcohol/drug use is acceptably consistent across the instruments. For example, the mean number of days of alcohol use at follow-up 3 for the ASI, the NIH-static and the NHI-dynamic assessments, respectively, is 5.5, 5.9, and 5.4 (see Figure 1); for days of heroin use at follow-up 3 the means are 0.9, 1.0, and 0.8 (see Figure 2).

The slight discrepancies (e.g. for cocaine use) may arise from differences in terminology among the instruments. Rigorously standardized instruments are desirable as they promote comparability. However, our analyses revealed several challenges in standardization. While the ASI is often considered a gold standard given its widespread use, we have found that a drug type such as amphetamine or cocaine may change in meaning depending on the drug epidemic development in the historical or geographic context. For example, amphetamine was popularly used first in the West Coast, but then methamphetamine became predominant in the West and Midwest; in some areas of the country, the pattern may differ. Similarly, the use of cocaine and crack cocaine has varied. Thus, the NHI has been adapted with slight changes to drug types and labels in order to accommodate the local and historical drug use context. In assessing reliability using different instruments or across different times, one needs to consider not only differences in instrument wording, but also the nature/form of the drug as it may change over time and the interpretation of the drug terminology within the regional context. In longitudinal research where measures are compared to assess changes over time, repeated assessments using the same instrument may not necessarily capture the same construct (i.e., the same drug), while retrospective recall over a long historical period (e.g. with the NHI-dynamic) has the advantage of specify the different types of drugs to reflect the pattern of drug epidemic developed in the area over time.

While the overall longitudinal patterns in our study appear reliable, as do levels of use and relationships between concurrent sources, there is somewhat less favorable consistency comparing either of the concurrent measures (ASI and NHI-static) with the longer-term retrospective recall (from NHI-dynamic). In particular, measures based on the NIH-dynamic form had greater discrepancy in the mean number of days using marijuana and generally lower correlations for all substances with ASI or the NHI-static collected at the follow-up 1. Possible explanations may include: retrospective recall may be limited by inaccurate memory, or participants are inclined to underreport current drug use. Relevant to possible inaccurate memory, the NHI-dynamic form administered at follow-up 2 required longer recall, approximately one year, back to the month corresponding to follow-up 1 than did the other instruments, which required recall of only the past month. The fact that the NHI-dynamic tends to provide higher levels of drug use lends some support for the second possible explanation, although memory loss still cannot be ruled out. While the ASI asks drug use in the past 30 days and the NHI-static form asks drug use in the past 4 weeks, we did not find differences between the two despite the slightly differently worded time frames.

The study results overall indicate that the general comparison of self-reported drug use across the ASI and NHI measures is very good. As Darke (1998) has noted, we "... should not make the mistake of expecting complete concordance of response as the gold standard for the acceptance of reliability of self-report" since memory is not perfect and some discrepancies are to be expected. The fact that in this study the patterns of drug use are

consistent across ASI and NHI assessment instruments supports their reliability for longitudinal examination of self-reported drug use.

Nevertheless, this study's limitations include reliance on self-report measures of substance use. However, self-report measures were corroborated by urinalysis for approximately half of the sample at each of the three study follow up points. In addition, the sample was selected from STD clinics, emergency departments and jails; therefore caution in generalizing results to samples of not-in-treatment drug users is warranted. However, because the overall sample is sufficiently large and diverse, and includes individuals who are often considered hidden populations and at high risk for drug abuse, characterizing the accuracy of self-reported substance use measures in this sample should provide useful information for measuring substance use in other populations of individuals who use alcohol and other drugs.

Longitudinal research is required for assessing changes. It has been suggested that the interval of repeated assessments should be based on the theory of change, so that the study can be designed efficiently to capture the change process and course. Nevertheless, in practice, the number of assessments and the interval of assessments are often influenced by logistical considerations (e.g., resources). It is encouraging that retrospective recalls such as those based on the NHI-dynamic instrument have revealed similar patterns of alcohol/drug use as those revealed by the ASI (which requires only very short-term recall for past 30 days). The NHI-dynamic data trajectories also have the added advantage of their additional detailed monthly information in between the data collection points represented by three administrations of the ASI, thus providing a viable option for collecting longitudinal research data and measuring changes.

Acknowledgments

This research was supported by Grant #P30 DA016383 from the National Institute of Drug Abuse to the second author.

Biographies

Debra A. Murphy, Ph.D., is a Research Psychologist and Director of the Health Risk Reduction Projects within the UCLA Integrated Substance Abuse Programs (ISAP). She has conducted studies primarily on HIV/AIDS behavioral research on children, adolescents, adults, and families, as well as other studies in the fields of substance abuse and orofacial injury.

Yih-Ing Hser, Ph.D., is Professor-in-Residence at UCLA ISAP and Director of the Center for Advancing Longitudinal Drug Abuse Research. She has been conducting research in the field of substance abuse and its treatment since 1980 and has extensive experience in research design and advanced statistical techniques applied to substance abuse data.

David Huang, Dr.P.H, is currently serving as senior statistician at UCLA ISAP. For the past ten years, he has been actively involved in large health education and epidemiological research projects at UCLA. He provides statistical support on several multi-center longitudinal surveys that provide health care to drug abuse patients in treatment institutes.

Mary-Lynn Brecht, Ph.D., is a Research Statistician and Director of Statistical Support for UCLA ISAP. She has had extensive training and experience in the development/adaptation, application, and integration of quantitative research methodologies, particularly those for

longitudinal analyses, with emphasis in the areas of drug-abuse, health systems, and treatment evaluation research.

Diane Herbeck, M.A., is a Research Associate and Project Director with UCLA ISAP. She has 10 years of mental health and substance abuse research experience, managing projects on such areas as psychotherapy process in alcohol treatment, the impact of maternal HIV on adolescents, and treatment program characteristics related to the use of evidence-based practices.

References

- Adair EBG, Craddock SG, Miller HG, Turner CF. Quality of treatment data: Reliability over time of self-reports given by clients in treatment for substance abuse. *Journal of Substance Abuse Treatment* 1996;13:145–150. [PubMed: 8880673]
- 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]
- Amsel Z, Mandell W, Mathias L, Hocherman I. Reliability and validity of self-reported illegal activities and drug use collected from narcotic addicts. *International Journal of the Addictions* 1976;11:325–336. [PubMed: 1262097]
- Anglin, MD.; McGlothlin, WH. Outcome of narcotic addict treatment in California. In: Tims, F.; Ludford, JP., editors. *Drug Abuse Treatment Evaluation: Strategies, Progress and Prospects*, NIDA Research Monograph 51, DHHS Publ. No. [ADM] 84–1329. Washington, DC: U.S. Government Printing Office; 1984. p. 106-128.
- Anglin MD, Hser Y, Chou C. Reliability and validity of retrospective behavioral self-report by narcotics addicts. *Evaluation Review* 1993;17:91–108.
- Appleby L, Dyson V, Altman E, Luchins DJ. Assessing substance use in multiproblem patients: reliability and validity of the Addiction Severity Index in a mental hospital population. *Journal of Nervous and Mental Disease* 1997;185:159–165. [PubMed: 9091597]
- Brecht, M-L. Los Angeles County drug trend update. Presented at the Community Epidemiological Workgroup Meeting; June, 2008; 2008 .
- Chou C, Hser Y, Anglin MD. Pattern reliability of narcotics addicts' self-reported data: A confirmatory assessment of construct validity and consistency. *Substance Use & Misuse* 1996;31:1189–1216. [PubMed: 8853237]
- Cook RF, Bernstein AD, Arrington TL, Andrews CM, Marshall GA. Methods for assessing drug use prevalence in the workplace: A comparison of self-report, urinalysis, and hair analysis. *International Journal of the Addictions* 1995;30:403–426. [PubMed: 7607776]
- Cooper GA, Allen DL, Scott KS, Oliver JS, Ditton J, Smith ID. Hair analysis: Self-reported use of “speed” and “ecstasy” compared with laboratory findings. *Journal of Forensic Sciences* 2000;45:400–406. [PubMed: 10782961]
- Croyle, RT.; Loftus, EF. *Improving episodic memory performance of survey respondents*. New York: Russell Sage Foundation; 1992.
- Darke S. Self-report among injecting drug users: A review. *Drug and Alcohol Dependence* 1998;51:253–263. [PubMed: 9787998]
- Darke S, Hall W, Wodak A, Heather N. Development and validation of a multi-dimensional instrument for assessing outcome of treatment among opiate users: The opiate treatment index. *British Journal of Addiction* 1992;87:733–742. [PubMed: 1591524]
- Ehrman RN, Robbins SJ. Reliability and validity of 6-month timeline reports of cocaine and heroin use in a methadone population. *Journal of Consulting and Clinical Psychology* 1994;62:843–850. [PubMed: 7962889]
- Fendrich M, Johnson TP, Sudman S, Wislar JS, Spiehler V. Validity of drug use reporting in a high-risk community sample: A comparison of cocaine and heroin survey reports with hair tests. *American Journal of Epidemiology* 1999;149:955–962. [PubMed: 10342805]
- Fendrich M, Vaughn CM. Diminished lifetime substance use over time: An inquiry into differential underreporting. *Public Opinion Quarterly* 1994;58:96–123.

- Hser Y, Anglin MD, Chou C. Reliability of retrospective self-report by narcotics addicts. *Psychological Assessment* 1992;4:207–213.
- Hser YI, Boyle K, Anglin MD. Drug use and correlates among sexually transmitted disease patients, emergency room patients, and arrestees. *Journal of Drug Issues* 1998;28:437–454.
- Johnson TP, Mott JA. The reliability of self-reported age of onset of tobacco, alcohol and illicit drug use. *Addiction* 2001;96:1187–1198. [PubMed: 11487424]
- Killeen TK, Brady KT, Gold PB, Tyson C, Simpson KN. Comparison of self-report versus agency records of service utilization in a community sample of individuals with alcohol use disorders. *Drug and Alcohol Dependence* 2004;73:141–147. [PubMed: 14725953]
- Kokkevi A, Richardson C, Palermou B, Leventakou V. Reliability of drug dependents' self-reports. *Drug and Alcohol Dependence* 1997;45:55–61. [PubMed: 9179507]
- Kuha, J. Measurement error in self-reported substance use. State College, PA: Penn State. The Methodology Center; 2001 .
- Lewandowski, CA.; Hill, TJ. A comparison of the Addiction Severity Index and Life History Calendar in assessing addiction and related factors among women: A methodological review. Poster presented at the meeting of the College for Problems on Drug Dependence; June 2003; Bal Harbour, FL. 2003.
- Liu S, Serdula MK, Byers T, Williamson DF, Mokdad AH, Flanders WD. Reliability of alcohol intake as recalled from 10 years in the past. *American Journal of Epidemiology* 1996;143:177–186. [PubMed: 8546119]
- Magura S, Freeman RC, Siddiqi Q, Lipton DS. The validity of hair analysis for detecting cocaine and heroin use among addicts. *The International Journal of the Addictions* 1992;27:51–69. [PubMed: 1537640]
- McGlothlin, WH.; Anglin, MD.; Wilson, BD. An evaluation of the California Civil Addict Program. DHEW Publication No. ADM 78–558. Washington, DC: U. S. Government Printing Office; 1977.
- McLellan AT, Luborsky L, Cacciola J, Griffith J, Evans F, Barr HL, O'Brien CP. New data from the Addiction Severity Index. Reliability and validity in three centers. *The Journal of Nervous and Mental Disease* 1985;173:412–423. [PubMed: 4009158]
- Mensch BS, Kandel DB. Underreporting of substance use in a national longitudinal youth cohort: Individual and interviewer effects. *Public Opinion Quarterly* 1988;52:100–124.
- Morrall AR, McCaffrey D, Iguchi MY. Hardcore drug users claim to be occasional users: Drug use frequency underreporting. *Drug and Alcohol Dependence* 2000;57:193–202. [PubMed: 10661670]
- Murphy DA, Durako S, Muenz LR, Wilson CM. Marijuana use among HIV+ and high-risk adolescents: A comparison of self-report through audio computer-assisted self-administered interviewing and urinalysis. *American Journal of Epidemiology* 2000;152:805–813. [PubMed: 11085391]
- Nunnally, JC. *Psychometric theory*. 2. New York: McGraw-Hill; 1978.
- Nurco, DN.; Bonito, AJ.; Lerner, M.; Balter, MB. The natural history of narcotic addiction: A first report. Washington, DC: Committee on Problems of Drug Dependence; 1975.
- Nyamathi A, Leake B, Longshore D, Gelberg L. Reliability of homeless women's reports: Concordance between hair assay and self report of cocaine use. *Nursing Research* 2001;50:165–171. [PubMed: 11393638]
- Parra GR, O'Neill SE, Sher KJ. Reliability of self-reported age of substance involvement onset. *Psychology of Addictive Behaviors* 2003;17:211–218. [PubMed: 14498815]
- Prusoff BA, Merikangas KR, Weissman MM. Lifetime prevalence and age of onset of psychiatric disorders: Recall 4 years later. *Journal of Psychiatric Research* 1988;22:107–117. [PubMed: 3261342]
- Shillington AM, Cottler LB, Mager DE, Compton WM. Self-report stability for substance use over 10 years: Data from the St. Louis Epidemiologic Catchment Study. *Drug and Alcohol Dependence* 1995;40:103–109. [PubMed: 8745131]
- Siddiqui O, Mott JA, Anderson TL, Flay BR. Characteristics of inconsistent respondents who have "ever used" drugs in a school-based sample. *Substance Use & Misuse* 1999;34:269–295. [PubMed: 10069753]

- Sobell LC, Kwan E, Sobell MB. Reliability of a drug history questionnaire DHQ. *Addictive Behaviors* 1995;20:233–241. [PubMed: 7484317]
- Sobell LC, Sobell MB, Leo GI, Cancilla A. Reliability of a timeline method: Assessing normal drinkers' reports of recent drinking and a comparative evaluation across several populations. *British Journal of Addiction* 1988;83:393–402. [PubMed: 3395719]
- Tulving, E. *Elements of episodic memory*. New York: Oxford University Press; 1983.

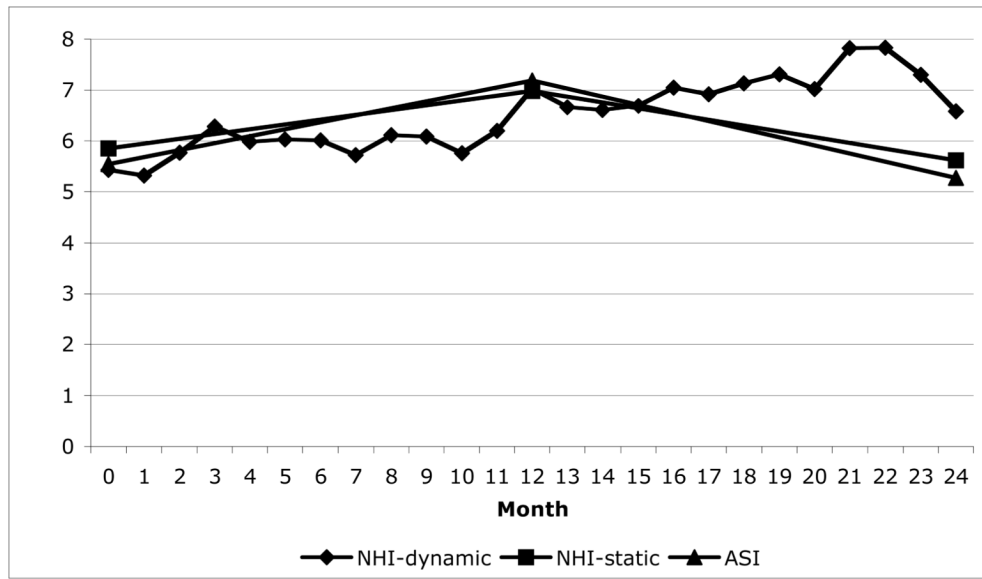


Figure 1.
Alcohol use for past month.

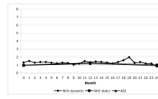


Figure 2.
Heroin use for past month.

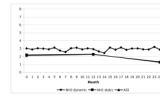


Figure 3.
Cocaine use for past month.

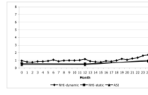


Figure 4.
Meth use for past month.

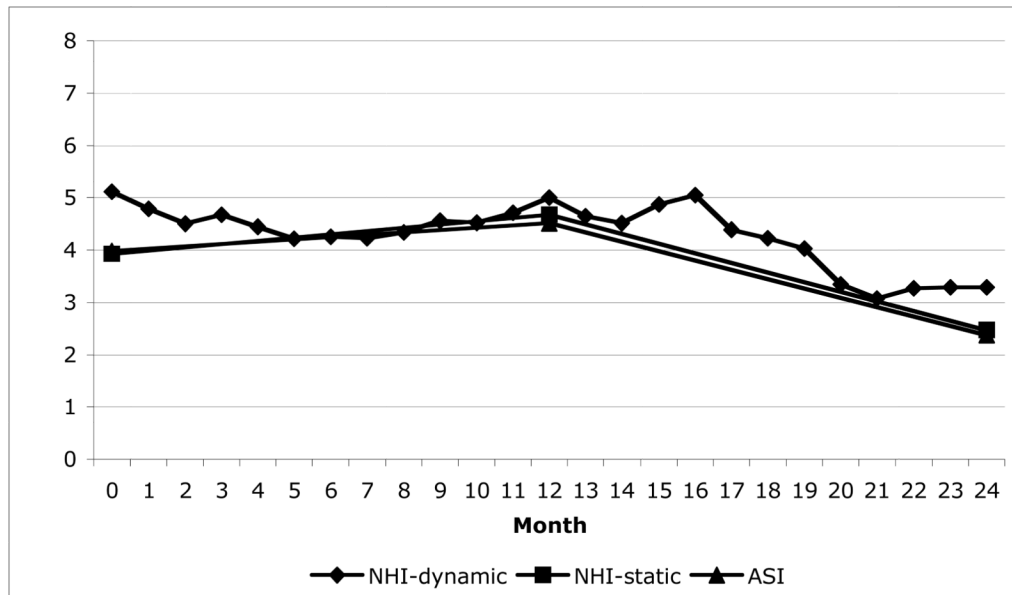


Figure 5.
Marijuana use for past month.

Table 1

Demographics of TUE Participants (N = 301) included in reliability analysis

Characteristic	%
Age ^a	
18–24	28.2
25–34	34.2
35–44	30.9
45+	6.7
Male	61.5
Ethnicity	
White	17.6
Black	53.8
Hispanic	25.9
Asian/Other	2.7
Marital Status	
Married	10.6
Single/never married	59.8
Divorce/widow	29.6
Education	
Less than high school	27.4
High school/GED	20.4
College	52.2
Employed	28.2
Primary drug	
Alcohol	11.5
Cocaine	41.0
Heroin	8.3
Marijuana	7.6
Methamphetamine	5.0
Other	6.8
None	19.8

^a*M* = 31.6; *SD* = 8.6.

Table 2

Correlations between measures from ASI, NHI-Static, and NHI-Dynamic of self-reported days of alcohol/drug use in the past month^a among users only.

	Correlations ^d (N)		
	ASI with NHI-Static	ASI with NHI-Dynamic	NHI-Static with NHI-Dynamic
Alcohol			
Follow-up 1	0.86 (178)	0.43 (178)	0.47 (189)
Follow-up 2	0.98 (137)	0.83 (137)	0.83 (144)
Follow-up 3	0.98 (164)	0.82 (163)	0.81 (164)
Heroin			
Follow-up 1	0.80 (22)	0.43 ^c (22)	0.52 (24)
Follow-up 2	0.99 (16)	0.89 (16)	0.89 (18)
Follow-up 3	0.99 (20)	0.72 (20)	0.77 (21)
Cocaine			
Follow-up 1	0.84 (93)	0.33 (93)	0.36 (97)
Follow-up 2	0.93 (56)	0.74 (56)	0.70 (61)
Follow-up 3	0.93 (62)	0.60 (62)	0.68 (63)
Methamphetamine			
Follow-up 1	0.90 (21)	0.25 ^b (22)	0.14 ^b (25)
Follow-up 2	0.65 (18)	0.70 (18)	0.93 (21)
Follow-up 3	0.97 (21)	0.95 (21)	0.97 (21)
Marijuana			
Follow-up 1	0.85 (130)	0.48 (130)	0.48 (136)
Follow-up 2	0.98 (93)	0.71 (94)	0.69 (101)
Follow-up 3	0.91 (107)	0.70 (108)	0.75 (108)

^a ASI reflects past 30 days; NHI-static past 4 weeks; and NHI-dynamic by month.

^b Correlation coefficients are NOT significantly different from zero;

^c Correlation coefficient is significantly different from zero at p=0.04; Otherwise,

^d All other correlation coefficients are significantly different from zero at p < .01