

# The Validity of Self-Reported Drug Use in Survey Research: An Overview and Critique of Research Methods

**Lana Harrison**

## ABSTRACT

Since illicit drug use is by definition illegal, the tasks of measuring incidence and prevalence and charting the course of the epidemic have fallen to survey researchers over the past 30 years. Although survey methods have obvious advantages over indirect measures such as arrests, seizures, and treatment admissions, they are frequently criticized because they rely on valid self-reporting of sensitive and highly stigmatized behavior. Validation studies conducted before the mid-1980s involving known samples of drug users or urinalysis techniques suggested that drug use was fairly accurately reported in self-report surveys. However, more recent validation studies conducted with criminal justice and former treatment clients using improved urinalysis techniques and hair analyses demonstrate that self-report methods miss a lot of recent drug use. A review of the research literature suggests that neither self-reports nor bioassays are wholly accurate, and both have inherent problems. However, because self-report measures are necessary to understand the complexity of causal and correlational attributes of drug abuse, it is necessary to determine what can be done to improve valid self-reporting. This chapter examines the research literature on validation studies to provide an overview of what is known about the accuracy of self-reported drug use.

## INTRODUCTION

How accurately can illicit drug use be measured in society? Drug use is an illegal activity and illicit drugs are illegal commodities; therefore, use cannot be measured by normal marketing procedures. Routinely compiled indicators such as police and court data on arrests and seizures, as well as clinical data on treatment admissions or drug-related medical emergencies provide a wealth of information, but can provide little information about the incidence and prevalence of use.

Surveys using self-report measures were initially developed as an alternative to anecdotal information or measures obtained from clinical, police, or court records. They were designed to assess the prevalence and frequency of illicit drug use in representative samples of the population. They were also designed to examine the correlates of illicit drug use and inform prevention and intervention efforts.

The sensitivity of collecting data on drug use has always made validity and reliability important issues. Survey research on drugs, where questions are asked about socially disapproved and illegal behaviors or socially marginal attitudes, may well generate inaccurate reporting and bias in survey estimates. Survey researchers recognize the need to design methods that elicit accurate and truthful reporting of drug use experience and attitudes. However, not much research has been conducted on the factors that improve an individual's reporting of sensitive information on questions about potentially embarrassing or self-incriminating behavior.

Even with their limitations, surveys are still a good measure of the nature and extent of drug use in a population, and provide information on the characteristics of drug users in a society. Treatment data provide very important information on the characteristics of people presenting themselves for treatment, but tell nothing about the characteristics of the pool of individuals from which those people are drawn or how those who enter treatment differ from those who do not. Likewise, drug users who become involved with the criminal justice system are not representative of drug users in general. Even ethnography, which frequently employs a loosely structured interview conducted in a more naturalistic setting, is restricted by its lack of generalizability to a known population, although it can provide a wealth of detailed information. Survey research can provide a more thorough profile of drug use and abuse among a broader cross-section of the population, and it can also provide a much greater range of information for use in designing intervention strategies. But the challenge is how to convince survey respondents to provide accurate information. Guarantees of anonymity and confidentiality are now standard fare in survey research on drugs. However, the research evidence suggests this is not enough to allay fears of some respondents in reporting recent drug use. Several of the articles in this monograph report on studies, particularly among criminal justice populations and treatment clients, in which recent drug use is underreported (see Mieczkowski and Newel, this volume; Wish et al., this volume; Magura and Kang, this volume).

## METHODS USED TO TEST THE VALIDITY OF SELF-REPORT

Validating self-report requires comparison to some method that is presumably more accurate. Over the past several decades, technologically sophisticated methods have been developed for analyzing drug metabolites in bodily fluids. Urine is most often used, but drugs have also been detected in blood, saliva, semen, meconium, perspiration, and hair (Smith and Liu 1986; Cone, this volume). Each biological specimen is unique and offers a somewhat different pattern of information regarding drug use over time. Also, each specimen has unique strengths and weaknesses regarding the type of information obtained from drug testing. The same testing methods are generally applied to the various bodily fluids. Testing methods fall into either screening or confirmations assays. Screening assays are generally valid—usually erring on the side of not identifying specimens that may contain drugs or their metabolites rather than identifying specimens as positive that do not contain drugs or drug metabolites. Confirmations assays are more expensive, but they are also more specific in identifying drug use.

Aside from biological assays, other methods have been explored for their potential to validate self-reports of drug use. Official record checks, such as checks of criminal justice or treatment records, have frequently been used to validate self-reports of drug use. Reports by family, close friends, or counselors have been used (Stephens 1972). Even polygraph tests have been used to validate self-reports of drug use (Clark and Tiffit 1966). These types of validation procedures, which rely on checking validity against external criteria, are examining external or empirical validity.

Internal validity procedures are performed in a cross-sectional survey to determine the amount of internal consistency across survey items. Internal validity checks may also be employed to assess internal consistency across survey items on repeated administrations of a longitudinal survey. For example, both the National Household Survey on Drug Abuse (NHSDA) (Substance Abuse and Mental Health Administration 1995*a*, 1995*b*) and the Monitoring the Future study (Johnston et al. 1995) of high school students demonstrate a high amount of internal consistency. In the Monitoring the Future survey, theoretically predicted relationships among a number of deviant behaviors have been demonstrated; estimates of friends' drug use closely parallel cumulative estimates of overall drug use (O'Malley et al. 1984). Analyses of NHSDA data show consistent patterns of self-

reported friendship with users of specific drugs, opportunity to use these drugs, and actual drug use (Harrell 1985). Analyses of the longitudinal followups of the Monitoring the Future data have also shown relationships among variables to persist over time (O'Malley et al. 1984). Drug use in the years following high school is highly consistent with and predictable from senior year drug use. Analyses also showed that past-year marijuana and alcohol use were more reliably measured than use in the past 30 days. Marijuana use was more reliably measured than the use of other illicit drugs (O'Malley et al. 1984).

### Urine Testing and Recent Challenges to Validity

There are obvious limitations to internal consistency checking and record checks, which is why the field has looked to bioassays to test the validity of self-report. Urine testing in particular has become more widespread and is considered to be quite valid, particularly with recent technological improvements. Earlier urine tests such as thin layer chromatography (TLC) have been found to be much less valid than the more recently developed tests such as enzyme multiplied immunoassay (EMIT) or fluorescence polarization immunoassay (FPIA). In a comparison of the three urine tests, the U.S. Justice Department found EMIT and FPIA to have false positive rates of about 0.2 to 2.5 percent (incorrectly identifying a negative specimen as positive), and false negative rates of 2.4 to 40.8 percent (incorrectly identifying a positive specimen as negative) (Visher 1991). The highest false negative rates were found for marijuana. Radioimmunoassay (RIA) was found to have a false positive rate of 0.1 to 4.1 percent, with the highest false positive rate (4.1 percent) associated with cocaine. The false negative rates for RIA ranged from 5.8 percent to 37.1 percent. The highest false negative rate was for marijuana. The validity of these urine tests was determined by comparing the EMIT and FPIA technologies to gas chromatography/mass spectrometry (GC/MS), which is presumed to be virtually 100 percent accurate. TLC showed a 0.3 to 3.1 percent false positive rate, but a false negative rate of 52 to 92 percent (Visher and McFadden 1991). All of the tests err on the side of not identifying a negative specimen as positive, which means they sacrifice the ability to correctly identify many specimens. However, the TLC test performance is significantly poorer than the other tests in terms of failing to identify positive specimens correctly. Since the late 1980s, several studies using the improved urinalysis technology have disputed the accuracy of self-report drug use surveys. It had generally been believed that estimates of self-reported drug use

were reasonably valid. In a review of self-report validation studies, Mieczkowski (1990) found that researchers reported validity rates of generally 70 percent or higher, and some even 90 percent or higher. The first large-scale study to cast doubt on the validity of self-reported drug use was the Drug Use Forecasting (DUF) study (Harrison 1989; National Institute of Justice 1990). Begun in several major U.S. cities in 1987, the study grew to include most of its current 23 sites by 1989. The DUF study employs urinalysis to measure drug use among those recently arrested and charged for serious crimes. Interviews are conducted in a central booking facility in several large U.S. cities, where privacy is not always available. Respondents are informed that the study is anonymous and confidential, and their participation will not have a bearing on their case. They are asked for a urine specimen at the end of the interview.

The DUF study has fairly consistently found that only about half of those who test positive for a drug report use in the past 2 to 3 days. Figure 1 compares drug use prevalence rates measured by urinalysis and self-report for the entire sample of arrestees participating in the DUF study in 1991. The most common way to interpret the congruence of urinalysis and self-report is to focus on just those with positive urinalyses and determine the percentage who accurately report their drug use. Of the 17.3 percent testing positive for marijuana, 9.3 percent report use in the past 3 days and 8.1 percent do not. Considering those who test positive, only about one-third to one-half admit their drug use. But this interpretation must be balanced against the interpretation that takes into account a fuller range of information. For example, notice that for marijuana 74.2 percent tested negative and (self-) reported no use of the drug in the past 3 days. Looking at the lower right hand corner of the marijuana grid, notice that 9.3 percent tested positive and admitted recent marijuana use. Therefore, in 83.5 percent of the cases, there was congruence between self-report and urinalysis. The rate of congruence for cocaine is 77.2 percent, and 95.3 percent for opiates. However, this measure of congruence is heavily influenced by the prevalence of drug use. The less likely the use, the higher the congruence rate.

Another way to look at the data is to compare the percentage who self-report use of a drug to the percentage who test positive for that drug. For example, 21.1 percent of the arrestees told the DUF interviewer that they

Marijuana				Cocaine									
Urinalysis				Urinalysis									
				Neg	Pos								
Self-Report	No	74.2%	8.1%	Self-Report	No	57.4%	21.6%						
	Yes	8.4%	9.3%		Yes	1.3%	19.8%						
								Opiates					
								Urinalysis					
								Neg	Pos				
Self-Report	No	91.6%	3.4%	Self-Report	No	91.6%	3.4%						
	Yes	1.2%	3.7%		Yes	1.2%	3.7%						

**FIGURE 1.** *Three-day self-report versus urinalysis among arrestees in major U.S. cities. Total DUF sample N = 34,720.*

SOURCE: National Institute of Justice 1990.

used cocaine in the past 3 days, but twice as many—41.3 percent—tested positive. About 4.9 percent reported recent opiate use, but 7.1 percent were found positive by urinalysis. The numbers are much closer in the case of marijuana, for which 17.7 percent admitted use in the past 3 days and 17.4 percent tested positive. The relative comparability between self-report and urinalysis for marijuana is largely impacted by those admitting use but not testing positive. Therefore, at least in the case of marijuana, self-report of use in the past 3 days appears to detect as much marijuana use as urinalysis—and this among individuals being interviewed in a jail setting. The fact that they underreport is probably not as surprising as the fact that many report validly. Despite assurances that they are participating in a confidential study, the respondents are interviewed in jail while awaiting arraignment; many are concerned about the outcome of their arraignment and anxious to talk to anyone. Perhaps it is not so surprising that recent cocaine and opiate use are not more validly reported among this population. Moreover, the congruence rate between self-report and urinalysis even among deviant groups in a less-than-ideal interviewing environment is noteworthy.

Another factor to consider is that earlier analyses also assume that the EMIT test is totally accurate. Recall that a Justice Department study found EMIT's false positive rate to fall between 2.1 and 2.5 percent for marijuana, opiates, and cocaine (Visher 1991). However, a more recent large scale study conducted by the National Institute on Drug Abuse (NIDA) found higher false positive rates in urine samples collected primarily in workplace settings, analyzed by EMIT, and for which positives were confirmed by GC/MS. Because the study involved laboratories that used GC/MS to verify samples that tested positive by EMIT, it can therefore only report on the false positive and negative rate of EMIT-screened positive specimens, and not on how well EMIT correctly identifies positive samples. In these tests, between 95 and 96 percent of the cocaine and marijuana samples found positive primarily by EMIT were confirmed by GC/MS, but only 53 to 55 percent of the opiates samples were confirmed (Stephenson 1992; Harrison 1995).

Urinalysis is not an exact science. EMIT can detect cocaine for 2 to 3 days in the urine. Opiates are detectable for 2 to 4 days, although detection time is generally limited to 2 days. Cannabis may be detected for up to 4 weeks. Further, the window of detectability is not a constant that applies in all cases, but is rather dependent upon the particular type of drug, the physical condition of the individual (i.e., state of hydration and fluid balance), the route of drug ingestion (i.e., intranasally, intra-venously), the amount of drugs used, and the individual's frequency of use (American Medical Association 1987).

Urinalysis has limitations in terms of what it can reveal about the validity of self-report. In addition, factors inherent to interviewing incarcerated people recently arrested for serious offenses limit the generalizability of the results to other samples. Validation studies employing urinalysis techniques have frequently been conducted on criminal justice populations, but the results are not generalizable to other populations because they oversample heavy drug users. Likewise, studies that use record checks to validate drug use may introduce bias simply because the characteristics of people likely to have records may differ significantly from those in the general population. Nevertheless such studies, and particularly the DUF study, have been used to call into question the validity of self-reported drug use in all surveys—regardless of the characteristics of the populations being surveyed. Findings on the validity of self-report from criminal justice populations have led to concern about the

validity of self-reported drug use from several policymakers (e.g., U.S. Senate 1990; General Accounting Office 1993).

## Hair Testing

Hair analysis is a newly developing technology being used as a check on the accuracy of self-reported drug use. It is increasingly being performed in numerous laboratories, some of which offer commercial drug testing services. Hair analysis is being used by private employers in pre-employment drug screening and tested in criminal justice settings. A pretrial diversionary program in New Orleans relies on hair testing, in addition to urinalysis and self-report, to monitor compliance with program rules of abstinence from illicit drugs (Mieczkowski et al. 1995). Hair testing has been receiving increased attention because of the less embarrassing circumstances of collection. Further, hair can be easily stored. Hair samples are generally collected from the vertex of the scalp, and then washed and dissolved with an acid or a strong base. With the hair sample prepared in this manner, the same types of analytical principles and technology used to analyze urine can be applied.

Although the technology of hair testing has progressed rapidly over the last decade, several highly controversial aspects of the procedure remain unresolved. It is still unclear how drugs enter the hair, creating concerns about contamination via exposure to cocaine dust particles, smoke, vapor, or drug solutions. At least two studies have found cocaine in the hair of children, suggesting that contamination is an important consideration (see Randall 1992; Smith et al. 1994). Cocaine has been shown to readily bind to hair, but binding depends on several physicochemical variables such as pH of exposure, ionic strength, and hair type (Kidwell and Blank 1994). Research shows much smaller amounts of cannabis are incorporated into the hair, and hair is not yet considered a good medium to test for cannabis use. Hair testing appears most valid for testing cocaine use (Mieczkowski and Newell, this volume; Hindin et al. 1994).

Another controversial issue in hair testing is the interpretation of dose and time relationships. Some research has suggested that the amount of drugs in the hair is proportional to the amount of use. Further, because hair grows at the rate of approximately  $1.5 \pm 0.15$  centimeters per month, it is believed that hair can be segmented to provide a record of an individual's drug use equivalent to the length of the hair. But studies with labeled cocaine have found only a limited dose and time relationship (Henderson et al. 1993; Cone 1994a).



Studies have shown that at any one time, about 85 percent of head hairs are growing (Hindin et al. 1994). There is also evidence of ethnic differences in hair test results; coarse, dark hair retains more of the drug than other hair types (Henderson et al. 1993; Cone 1994*b*; Kidwell and Blank 1994). Despite these limitations, hair is increasingly being used to detect drug use. Several studies comparing hair, urine, and self-report results are reported in chapters in this monograph. These studies suggest that hair analyses disclose more recent drug users than can be found through either urinalysis or self-reports.

One of the early studies comparing hair, urine, and self-report was conducted by Mieczkowski and colleagues (1991*b*) in Florida using a prototype of the DUF study. Hair was analyzed by RIA, and both EMIT and FPIA were used to test urine. Mieczkowski and colleagues concluded that about four times as many arrestees had a positive hair assay as self-reported cocaine use within the previous 30 days. There was a ninefold increase in the number who had hair positive for opiates as compared to self-reported opiate use in the past 30 days (Mieczkowski et al. 1991*b*). Mieczkowski and colleagues also found that individuals were less likely to accurately report use in the immediate past (48 hours) and more likely to report use over longer time periods (30 and 60 days). They determined that self-report was least reliable for cocaine (Mieczkowski et al. 1991*b*). There were many inconsistencies in comparisons of the urine and hair samples, which to some extent was expected because the hair was analyzed for the past 60 days. However, not all the inconsistencies can be explained by the differences in the time frames of the tests employed. In fact, in a study of probationers in which an average of 5.9 urine samples were obtained per month, only 46 percent of positive RIA for hair (RIAH) tests were confirmed by urinalysis for cocaine and only 60 percent for heroin (Baumgartner et al. 1989; also reported in Hindin et al. 1994).

Some discrepancies between urine and hair analysis results cannot easily be explained. Recall that RIA, which has been extensively employed in analyzing hair samples, was found in the U.S. Justice Department studies mentioned earlier to have a 4.1 percent false positive rate for cocaine based on analysis of urine. Nevertheless, researchers tend to conclude that hair analysis is most accurate in detecting cocaine use. The host of unresolved issues surrounding hair analyses give reason for concern in attempting to validate drug use. Most studies that have been conducted analyzing drugs in hair have used relatively small numbers of subjects and have failed to include

proper controls. Washing may be problematic because drugs may be removed by this procedure. The effects of shampooing and cosmetic treatments such as dyeing, perming, or bleaching, and exposure to ultraviolet light or other external contaminants may alter the presence of the drug in the hair shaft (Harkey and Henderson 1989; Henderson et al. 1993). Research has shown that uptake and washout rates of cocaine in hair, for example, vary extensively between individuals and may be related in part to differences in hygiene. It is critical that validation of self-report data by drug testing be performed with careful consideration of the limitations imposed by the testing methodology and the biological specimen.

### Validation Studies

There are studies in the literature that suggest relatively high rates of self-reported drug use. For example, Zanis and associates (1994) found that for a sample of patients in methadone treatment for at least 6 months, 13 percent and 19 percent of those testing positive by EMIT for opiates and cocaine respectively, failed to self-report use in the previous month. Additionally, 58 percent and 28 percent of those with negative urines for opiates and cocaine, respectively, reported use of the drug during the previous month. The results of urinalyses from 154 subjects in four cities in a study of human immunodeficiency virus (HIV) risk behaviors showed that 71.2 percent tested positive, and 73.2 percent reported using cocaine in the past 48 hours. This was a highly drug-involved sample; 76 percent reported injecting drugs in the past 30 days. Self-reports and urinalysis results agreed for 85 percent of the heroin users. Self-reported drug use in the past 48 hours was not confirmed by urinalysis among 9.7 percent of those reporting heroin use and 7.8 percent of those reporting cocaine use. Positive urinalysis results were found in 5.2 percent and 5.8 percent, respectively, of respondents who did not self-report heroin or cocaine use in the past 48 hours (Weatherby et al. 1994). Therefore, self-report produced much higher rates of drug use than obtained by urinalysis. Likewise, analyses by Weatherby and colleagues (1994) suggest that heavily drug-involved individuals can self-report recent drug use fairly validly.

In a validity study conducted among a workplace population, which more closely resembles a general population than treatment or criminal justice populations, self-reports were found to quite reliably measure drug use. Cook and colleagues (this volume) found that self-reports produced higher prevalence rates than either urinalysis or hair analysis. For the entire sample (N = 800) of employees from a large

steel plant, only 7.8 percent tested positive for any drug by urinalysis, while 9.4 percent reported recent drug use. For the subsample that had hair tests, 6.2 percent were positive for an illicit drug and 9.9 percent reported recent use. The most frequently detected drug was marijuana, and little cocaine use was found by testing. It was also anticipated that some of the subjects who tested positive for morphine or sedatives failed to report prescription drug use. Another important caveat is that the study used much lower cutoff levels than recommended by NIDA for determining a urine specimen as positive (because the analyses were being conducted for research purposes only). Because of the small number with positive drug assays despite the lowered cutoff levels, the validity analyses were combined across all drug types. Although self-report methods produced higher prevalence rates than testing, Cook and colleagues found only about half of those positive for any drug by either urine or hair test self-reported recent use. They concluded that it is necessary to use multiple assessment methods to estimate self-report. Cook and colleagues also varied the method of data-collection setting between telephone interview, personal interview in the workplace, group interview in the workplace, and personal interview offsite. They found rates of drug use self-report were highest in the individual workplace interview and lowest in the group interview in the workplace. The results nevertheless suggest that self-report methods appear to provide good measures of prevalence—in this study, higher than those generated by the assay tests.

It seems clear that although drug use may vary substantially among different populations such as household members, students, and arrestees, the accuracy of their self-report may also vary substantially. The research literature suggests that self-report may be the least reliable among criminal justice clients. Magura and Kang (this volume) report the results of two validity studies conducted using similar methods in New York City. One study included a sample of patients in two methadone treatment programs recruited because clinic records showed they had tested positive by urinalysis for cocaine. The other study included a sample of criminally involved young adults. The young adults were recruited while they were in jail, but were followed up in the community about 5 months after release. Self-report information and both urine and hair samples were obtained on all the clients at followup interviews. For the methadone sample, 60 percent self-reported recent cocaine use and 80 percent were positive by RIAH. For the young adult sample, 23 percent self-reported recent cocaine use, but almost three times as many—67 percent—were positive by RIAH. This led the researchers to

conclude that self-report is more valid for treatment clients than for criminal justice clients. An interesting aside is that, whereas only 23 percent of the criminal justice-involved young adults self-reported recent cocaine use, 75 percent reported recent marijuana use, and fully 41 percent reported drug dealing in the past month. This would indicate that the young adults were not afraid of self-reporting sensitive information. Magura and Kang suggest that for the young adults, cocaine (or more specifically, crack) had become stigmatized, even though dealing of these drugs was not; while the young adults reported drug dealing, they were more reluctant to report cocaine use. There may also be an important explanation overlooked by the authors, which is that hair analysis might detect cocaine that had entered the young adults' hair through environmental contamination via the youth's handling of cocaine/crack for sale.

With respect to the validity of self-report among treatment clients, the research literature suggests that self-report is more accurate at intake. That is, clients are more likely to provide self-reports that are congruent with urine or hair test results in the beginning stages of treatment than they are at followup posttreatment (Wish, this volume; Hindin et al. 1994). For example, Hindin and colleagues (1994) found that among 109 entrants to two New England treatment facilities, 89 percent of the 87 found positive for cocaine by hair and 96 percent of 45 heroin positives were confirmed by self-report. However, among the 86 followed up, only 51 percent of the 43 found positive for cocaine by hair and 67 percent of 18 heroin positives were confirmed by self-report. This is an important finding because it suggests the importance of validating self-report in studies measuring treatment outcome.

### Limitations of Validity Research

Validity research is still in its early stages. It is hampered by the limitations of technology, but also by the lack of sophisticated knowledge about critical elements conducive to the honest reporting of sensitive information in a survey environment. There are undoubtedly multiple influences on respondents in terms of their ability and desire to provide a valid response. These factors include setting, real or perceived consequences of reporting use, literacy, clarity of questions, and memory. Studies validating self-report have frequently not taken advantage of procedures to maximize accurate self-reporting of illicit drug use. Although studies typically promise anonymity and confidentiality, confidentiality can be compromised

by administration procedures that require respondents to provide their answers aloud to interviewers.

In a series of methodological studies undertaken in conjunction with the NHSDA, significantly higher rates of drug use were found using self-administered answer sheets as opposed to having respondents answer aloud to interviewers. The methodological field test found respondents were 1.6 times more likely to report cocaine use in the past year and 2.4 times more likely to report use in the past 30 days on the self-administered answer sheets. The increase in reported drug use was a function of the recency of the event, with few differences in lifetime rates, more difference in past-year rates, and the greatest difference in rates of past month drug use. A more recent national field test of even the cigarette questions revealed significantly higher rates of cigarette use reported using a self-administered as opposed to an interviewer-administered answer sheet (SAMHSA 1995*a*).

Similar findings about the impact of self- versus interviewer-administered questions on drug use have been found in the National Longitudinal Survey of Labor Market Experience, Youth Cohort (NLSY). Substantially less cocaine and marijuana use were reported in the interviewer-administered conditions. Respondents were more likely to report marijuana as opposed to cocaine use (Schober et al. 1992). Mensch and Kandel (1988) also found self-administered answer sheets to yield higher reports of drug use than interviewer-administered questions. Likewise, a study of prison inmates found more reports of drug use using self-administered versus interviewer-administered questionnaires (McElrath 1994).

Another finding from the series of methodological studies conducted in conjunction with the NHSDA that has been replicated in a number of studies: The more stigmatized the drug, the more prevalence rates are suppressed. Marijuana use is reported more validly than cocaine use. This finding has been replicated in several studies (Harrison 1992, 1995; Fendrich and Xu 1994; Mieczkowski et al. 1991*a*). Currently the most stigmatized drug appears to be cocaine (or more specifically, crack). In the DUF study, comparing the percentages who self-report use of the respective drugs to the percentages who test positive leads to the conclusion that arrestees are most willing to admit marijuana use, followed by opiates, amphetamines, and then cocaine (Harrison 1992).

To maximize reporting the use of stigmatized drugs, it is vital to use procedures that maximize confidentiality. This includes using self-administered rather than speak-aloud interviews. Most of the research that has been done has involved paper-and-pencil questionnaires, but research is also beginning on the impact of computer-assisted self-interviewing (CASI) in improving the validity of self-reported drug use (Lessler and O'Reilly, this volume). The setting also needs to be explored in validity research, because it logically makes sense that recent arrestees interviewed in jail prior to being sentenced may not provide especially valid reports of recent drug use. There is also very limited research on respondents' perceptions of risk of providing truthful information to sensitive drug questions (Willis et al., this volume). The scientific community needs to engage in a systematic research program, varying different aspects of the interview environment and methodology to improve understanding of factors that can improve the validity of self-reports. Further, other factors that contribute to validity in survey research must not be forgotten (Gfroerer et al., this volume). The generalizability of survey research is predicated on proper selection procedures from a known universe. Also, surveys must have a respectable response rate to ensure that bias is not introduced if members of the target population are consistently underrepresented in the sample.

## DISCUSSION

Concern is often expressed about the validity of survey data on drug use. Survey-generated estimates of drug use are frequently criticized on the grounds that many survey respondents are not honest in reporting illicit drug use. At this point, it is not possible to judge how validly individuals report their drug use in surveys. There are inherent difficulties in trying to measure the validity of self-report based on the current available methods. The available chemical test(s) and methods now used to judge the validity of self-reported data on drug use surveys have limitations. Urine tests have a narrow window of detectability, which greatly reduces their usefulness. Studies based on record checks may be biased simply because the characteristics of people likely to have records may differ significantly from those in the general population.

Perhaps hair analysis holds the greatest promise in providing a standardized external validity criterion measure because it measures drug use over a longer period of time and hair samples can be obtained

unobtrusively. However, more research is needed before this method can be deemed reliable and valid. Research has yet to answer unresolved problems with hair testing. The consensus of scientific opinion is that hair analysis for the presence of drugs of abuse is unreliable and is not generally recognized by qualified experts as effective. In October 1994, the Society of Forensic Toxicologists held a meeting sponsored by NIDA to review the available research on hair testing. The participants concluded there were even more questions about hair testing for drugs of abuse than before (Society of Forensic Toxicologists 1994). Therefore, hair analysis cannot currently be described as useable with acceptable accuracy.

Thus far, the largest problem with most external criteria validation studies is that results are inherently not generalizable. Urinalysis studies, and now studies using hair analysis, are most often conducted on populations that have much higher rates of drug use than the general population. Respondents have not been randomly recruited from some definable population, so results can be generalized only to a discrete population group. Notably, validity rates for criminal justice populations cannot be generalized to the general population. Criminal justice populations may be less honest because they could be heavily penalized if their drug use were known to authorities. Consequently, those involved with the criminal justice system may make different decisions about candor in interviews than would the general population.

The research literature suggests that the validity of self-report varies by population subgroup. For example, arrestees are much less likely to provide honest reports of recent drug use than people in treatment. Self-report surveys of employees found higher prevalence rates based on self-report than on urine analysis or hair analysis. However, there is still a dearth of good studies that look at the validity of self-report in general populations. Most research conducted on validating self-report has focused on criminal justice and treatment populations, and is limited in its ability to determine how accurately respondents report drug use in general population surveys (such as household and student surveys).

Despite the concerns with the generalizability of the results of most validation studies, research does point to some general conclusions that may be drawn about the validity of self-reported drug use in a survey environment. Clues are provided in the consistency of results across several studies that show differences in self-reporting by drug type. The pattern of reporting is consistent with the social

desirability hypothesis about more stigmatized drugs such as cocaine, the least validly reported. Also, as the use of drugs becomes more recent, it appears to be subject to increasing bias; respondents are most willing to report lifetime use and least willing to report use that occurred in the very recent past. This has further implications for the usefulness of urinalysis to validate self-report information such as that derived from household or student surveys, since they are generally concerned with measures of lifetime, past-year, and past-month drug use, not use in the past 2 to 3 days (which urinalysis is able to measure). Another finding is that the use of self-administered questionnaires tends to produce higher prevalence rates (and ostensibly, more valid data) than interviews in which the respondents must speak their responses aloud.

Some surveys undoubtedly obtain more valid information than others. Even within surveys, differences in interviewer styles and presentation influence validity. Probably what is most amazing is that individuals will admit to illicit drug use in surveys. There are definite limitations to survey research on drug use, but perhaps many of those can be overcome with research designed to further improve the validity of self-report. It is important to remember that most validity research, in fact, shows quite high congruence rates between self-report and assay results. Factoring in that some of the differences between self-report and urinalysis found in validation studies are also due to the interview process (i.e., question wording, interview expectations, setting) leads to the conclusion that even among at-risk populations, there is a high degree of congruence between self-report and urinalysis. Of course, the research also shows a lot of individual variation; many validation studies find only about half of those testing positive for an illicit drug report using that drug. But do not forget the limitations of the testing technology. The current state of the science suggests that the most appropriate presentation of results from hair testing, in particular, would be in the aggregate, and not at the individual level.

In conclusion, self-report information is always going to be necessary, because biological assays can only corroborate drug use. Assays cannot determine the age at which individuals initiated drug use, individuals' attitudes about the risk of harm, perceptions of drug availability, other factors that may co-vary with drug involvement (such as other deviant behavior), and even whether individuals have received treatment in the past. It will always be necessary to rely on self-report to collect some sensitive information, which suggests it is imperative to conduct research on those factors that can be



manipulated within a survey environment to increase the validity of self-reporting of sensitive information. Since much of the research that scientists, policymakers, the media, and other interested individuals use to inform themselves about drugs is based on self-report, it is important to engage in more systematic and rigorous scientific studies to improve the validity of self-report. To really determine how accurately self-report survey research methods measure drug use, it is essential to gather scientific data on what methodological or environmental circumstances can be manipulated to improve the validity of self-report for what types of population subgroups.

## REFERENCES

- American Medical Association, Council on Scientific Affairs. Scientific issues in drug testing. *JAMA* 257(22):3110-3114, 1987.
- Baumgartner, W.A.; Hill, V.A.; and Bland, W.H. Hair analysis for drugs of abuse. *J Forensic Sci* 34(76):1433-1453, 1989.
- Clark, J.P., and Tiffit, L.L. Polygraph and interview validation of self-reported deviant behavior. *Am Sociol Rev* 31:516-523, 1966.
- Cone, E. "Drug Testing in Hair." Paper presented at the Society of Forensic Toxicologists (SOFT) Conference on Drug Testing in Hair, Tampa, FL, Oct. 29-30, 1994a.
- Cone, E. "Potential for Ethnic Bias in Hair Testing." Paper presented at the SOFT Conference on Drug Testing in Hair, Tampa, FL, Oct. 29-30, 1994b.
- Fendrich, M., and Xu, Y. The validity of drug use reports from juvenile arrestees. *Int J Addict* 29(8):971-985, 1994.
- General Accounting Office (GAO). *Drug Use Measurement Strengths, Limitations, and Recommendations for Improvement*. Report GAO-PEMD 93-18. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1993.
- Harkey, M.R., and Henderson, G.L. Hair analysis for drugs of abuse. In: Basalt, R., ed. *Advances in Analytical Toxicology*. Chicago: Year Book Medical Publishers, 1989.
- Harrell, A.V. Validation of self-report: The research record. In: Rouse, B.; Kozel, N.; and Richards, L., eds. *Self-Report Methods of Estimating Drug Use*. National Institute on Drug Abuse Research Monograph 57. DHHS Pub. No. (ADM)85-1402. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 12-21.

- Harrison, L.D. "The Validity of Self-Reported Drug Use Among Arrestees." Paper presented at the 41st meeting of the American Society of Criminology, Reno, Nov. 8-12, 1989.
- Harrison, L.D. The validity of self-reported data on drug use. *J Drug Issues* 25(1):91-111, 1995.
- Harrison, L.D. Trends in illicit drug use in the USA; Conflicting results from national surveys. *Int J Addict* 27(7):817-847, 1992.
- Henderson, G.L.; Harkey, M.R.; and Jones, R. "Hair Analysis for Drugs of Abuse." Final report of activities under National Institute of Justice/ National Institute on Drug Abuse grant no. NIJ90-NIJ-CX-0012, 1993.
- Hindin, R.; McCusker, J.; Vickers-Lahti, M.; Bigelow, C.; Garfield, F.; and Lewis, L. Radioimmunoassay of hair for determination of cocaine, heroin, and marijuana exposure: Comparison with self-report. *Int J Addict* 29(6):771-789, 1994.
- Johnston, L.D.; O'Malley, P.N.; and Bachman, J.G. *National Survey Results on Drug Use from the Monitoring the Future Study, 1975-1992*. Rockville, MD: National Institute on Drug Abuse, 1995.
- Kidwell, D.A., and Blank, D.L. "Incorporation of Drugs of Abuse into and Their Removal from Hair." Paper presented at the SOFT Conference on Drug Testing in Hair, Tampa, FL, Oct. 29-30, 1994.
- McElrath, K. A comparison of two methods for examining inmates' self-reported drug use. *Int J Addict* 29(4):517-524, 1994.
- Mensch, B., and Kandel, D. Underreporting of substance use in a national longitudinal youth cohort: Individual and interviewer effects. *Pub Opin Q* 52:100-124, 1988.
- Mieczkowski, T. The accuracy of self-reported drug use: An evaluation and analysis of new data, In: Weisheit. R., ed. *Drugs and Crime and the Criminal Justice System*. Cincinnati: Anderson Publishing Co., 1990.
- Mieczkowski, T.; Barzelay, D.; Gropper, B.; and Wish, E. Concordance of three measures of cocaine use in an arrestee population: Hair, urine and self-report. *J Psychoactive Drugs* 23(3):241-249, 1991a.
- Mieczkowski, T.; Landress, H.J.; Newel, R.; and Coletti, S.D. "The Concordance of Drug Use Indicators: Urine, Hair and Self-Report in an Arrestee Population." Report of activities under National Institute of Justice contract no. 90-IJ-CX-0023, 1991b.

- Mieczkowski, T.; Mumm, R.; and Connick, H.F. The use of hair analysis in a pretrial diversion program in New Orleans. *Int J Offender Ther* 39(3):222-241, 1995.
- National Institute of Justice. *DUF: 1988 Drug Use Forecasting Annual Report*. Washington, DC: National Institute of Justice, 1990.
- O'Malley, P.M.; Bachman, J.G.; and Johnston, L.D. Reliability and consistency in self-reports of drug use. *Int J Addict* 18:805-824, 1984.
- Randall, T. Infants, children test positive for cocaine after exposure to second-hand crack smoke. [Letter to the editor]. *JAMA* 267(8), 1992.
- Schober, S.E.; Caces, M.F.; Pergamit, M.R.; and Branden, L. Effect of mode of administration on reporting of drug use in the National Longitudinal Survey. In: Turner, C.F.; Lessler, J.T.; and Gfroerer, J.C., eds. *Survey Measurement of Drug Use: Methodological Studies*. DHHS Pub. No. (ADM)92-1929. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992. pp. 267-276.
- Smith, F., and Liu, R. Detection of phenobarbital in bloodstains, semen, seminal stains, perspiration stains and hair. *J For Sci* 26:582-586, 1986.
- Smith, F.P.; Kidwell, D.A.; and Cook, L.F. "Cocaine in Children's Hair When They Live with Drug Dependent Adults." Paper presented at the SOFT Conference on Drug Testing in Hair, Tampa, FL, Oct. 29-30, 1994.
- Society of Forensic Toxicologists (SOFT). *Consensus Opinion Summarizing the Current Applicability of Hair Analysis to Testing for Drugs of Abuse*. SOFT, June 1990.
- Stephens, R.C. The truthfulness of addict respondents in research projects. *Int J Addict* 7:549-558, 1972.
- Stephenson, R.L. "The Potential for Collision: Drug Testing in the Workplace and in the Criminal Justice System." Paper presented at the National Institute of Justice/Bureau of Justice Assistance Third Annual Evaluating Drug Control Initiatives Conference, Washington, DC, July 27, 1992.
- Substance Abuse and Mental Health Services Administration (SAMHSA). *National Household Survey on Drug Abuse: Main Findings 1992*. DHHS Pub. No. (SMA)94-3012. Rockville, MD: SAMHSA, 1995a.
- Substance Abuse and Mental Health Services Administration (SAMHSA). *Preliminary Estimates from the 1994 National Household Survey on Drug Abuse*. Rockville, MD: SAMHSA, 1995b.

- U.S. Senate. *Drug Use in America: Is the Epidemic Really Over?* Summary of findings of a majority staff study prepared for the Judiciary Committee, 101st Congress, Dec. 19, 1990.
- Visher, C. A comparison of urinalysis technologies for drug testing in criminal justice. *National Institute of Justice Research Report*. Washington, DC: U.S. Department of Justice, National Institute of Justice, 1991.
- Visher, C., and McFadden, K. A comparison of urinalysis technologies for drug testing in criminal justice. *National Institute of Justice Research in Action*. Washington, DC: U.S. Department of Justice, National Institute of Justice, 1991.
- Weatherby, N.L.; Needle, R.; Cesari, H.; Booth, R.; McCoy, C.B.; Watters, J.K.; Williams, M.; and Chitwood, D.D. Validity of self-reported drug use among injection drug users and crack cocaine users recruited through street outreach. *Eval Prog Plan* 17(4):347-355, 1994.
- Zanis, D.A.; McLellan, A.T.; and Randall, M. Can you trust patient self-reports of drug use during treatment. *Drug Alcohol Depend* 35:127-132, 1994.

#### AUTHOR

Lana Harrison, Ph.D.  
Associate Director  
Center for Drug and Alcohol Studies  
University of Delaware  
77 East Main Street  
Newark, DE 19716

**[Click here to go to page 37](#)**