

Review Article

Sources of Error in Substance Use Prevalence Surveys

Timothy P. Johnson

Survey Research Laboratory, University of Illinois at Chicago, 412 S. Peoria Street, Chicago, IL 60607, USA

Correspondence should be addressed to Timothy P. Johnson; timj@uic.edu

Received 30 April 2014; Accepted 13 October 2014; Published 5 November 2014

Academic Editor: David A. Sleet

Copyright © 2014 Timothy P. Johnson. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Population-based estimates of substance use patterns have been regularly reported now for several decades. Concerns with the quality of the survey methodologies employed to produce those estimates date back almost as far. Those concerns have led to a considerable body of research specifically focused on understanding the nature and consequences of survey-based errors in substance use epidemiology. This paper reviews and summarizes that empirical research by organizing it within a total survey error model framework that considers multiple types of representation and measurement errors. Gaps in our knowledge of error sources in substance use surveys and areas needing future research are also identified.

1. Introduction

For the past 50 years, social and epidemiologic surveys have been employed to estimate and track the substance use patterns of representative samples of both adolescents and adults in the United States and other countries. Although many of these surveys are of exceptional quality and rigor (e.g., the Monitoring the Future Survey, the National Survey of Drug Use and Health, and The Youth Risk Behavior Survey), for almost as long, there have been methodological criticisms and skepticism regarding their ability to accurately portray the behaviors they seek to measure [1–7]. Addressing these questions is important given the lack of alternative methodologies for efficiently monitoring substance use behavior within large national and subnational populations. The goal of this paper is to review and summarize the available empirical evidence addressing these questions, to identify gaps in our knowledge base regarding this issue, and to make some recommendations for needed future research to address those knowledge gaps.

2. The Total Survey Error Model

A useful framework for conceptualizing error in substance use surveys is the total survey error (TSE) model. The TSE model first delineated by Groves [8] focused on sampling, coverage, nonresponse, and measurement errors in surveys.

This model successfully organized decades of empirical research within a single unifying theoretical framework. An expanded elaboration of the TSE model has been more recently presented by Lavrakas [9], in which he identifies two general classes of errors, measurement and representation, and then explores multiple subclasses of errors within each. Table 1 lists the various elements of the Lavrakas TSE model.

Briefly, errors of representation are those concerned with technical problems that may impede a survey's ability to accurately mirror the population that the survey seeks to represent. These include failure to use sample frames that provide adequate coverage of the population being studied (coverage errors), imprecision in the sample(s) drawn from a sample frame (sampling error), errors associated with failure to contact or complete interviews with all sampled respondents, and failure to obtain answers to all questions included in a survey instrument (nonresponse errors), as well as failure to make adequate adjustments for complex sample designs and survey nonresponse (adjustment errors).

In contrast, errors of measurement involve failures to adequately assess the variables of interest in a survey. These include specification errors, which involve failures to correctly conceptualize survey constructs, and measurement errors, which include factors external to the construct being measured that nonetheless influence measurement quality. Processing errors are defects in the construction of survey data sets and/or final analytic variables and inferential errors

which involve difficulties or failures when making adequate sense of the final survey data. The following two sections organize the empirical literature concerned with errors in substance use and misuse surveys within this TSE framework. Each of these error sources, of course, is broadly relevant to health survey research in general. Our goal here is to review their relevance specifically to substance use surveys.

3. Errors of Representation

3.1. Coverage Errors. Errors in coverage are generally a consequence of employing a survey sampling frame that does not include all individuals in the population being studied, or, alternatively, by employing methods that do not provide all members of the population of interest some probability of being sampled. As with all other elements of the TSE framework, this type of error is not unique to substance use surveys. Nonetheless, because likelihood of falling into a potential sample frame may in some cases be associated with substance use behaviors, substance use research may be uniquely vulnerable to coverage error.

In most community epidemiological surveys, there are many social groups that may be systematically excluded from commonly available sample frames. Some of these groups include homeless persons, individuals currently hospitalized, college students living in dormitories, persons incarcerated in the criminal justice system, and members of the military living on military bases. Substance use may be particularly high within some of these nonresidential populations [10, 11]. Weisner and colleagues [7] investigated this problem by comparing prevalence estimates from a general population community survey with data obtained from interviews with nonhousehold populations found in several inpatient and outpatient settings, such as alcohol, drug or mental health treatment, criminal justice, and/or welfare services. Not surprisingly, substance use was much more common among persons in these settings. For example, 11.3% of the household sample was defined as problem drinkers, compared to 43.1% of those found in nonhousehold agency settings. The disparities were even greater for indicators of weekly drug use (5.5% in the household sample versus 36.5% in the agency sample) and both problem drinking and weekly drug use combined (2.2% in household sample versus 18.7% in agency sample). Other research provides similar evidence of increased substance use and misuse among persons less likely to be sampled within single family households as part of community-based epidemiologic surveys [12].

There is also some evidence in the US that failure to incorporate cell-phone-only households into random digit dialed (RDD) telephone samples can lead to underrepresentation of young adults who are at higher risk for substance use behaviors. Delnevo et al. [13] found significantly decreased measures of binge drinking and heavy alcohol consumption between 2001–2003 and 2003–2005 in the national Behavioral Risk Factor Surveillance System (BRFSS) telephone surveys. Other research, employing the US National Health Interview Survey, which relies on face-to-face interviews, has demonstrated that adults in cell-phone-only households are more likely to report past year binge drinking behavior (37.6%),

TABLE 1: Elements of the Lavrakas (2013) total survey error model.

Errors of representation	Errors of measurement
Coverage errors	Specification errors
Sampling errors	Measurement errors
Nonresponse errors	Processing errors
Adjustment errors	Inferential errors

compared to those residing in households with landlines (18.0%), and to those in households with no telephone service (23.0%; Blumberg et al. [14]). The effects of excluding cell phone-only households from survey estimates of binge drinking are particularly serious for young adults (aged 18–29 years) and low income persons [15]. Similar relationships between type of phone subscribership and substance use reports have been identified in Australia [16] and in other US studies [17]. As rates of cell-phone-only residences continue to grow [18], the coverage error associated with excluding them from telephone samples will only increase, and it will become increasingly difficult to produce credible prevalence estimates using traditional landline-only sample frames.

School-based surveys are also subject to coverage errors, as substance use rates have been shown to be higher among adolescents who drop out of school [19, 20]. Hence, surveys of adolescents that are school based often underestimate substance use within this population, although it is important to acknowledge that many school-based surveys attempt to make no generalizations to nonschool populations. A recent analysis by Gfroerer and colleagues [21] using pooled data from the 2002–2008 NSDUH (National Survey of Drug Use and Health, previously known as the National Household Survey of Drug Abuse or NHSDA) surveys reported that substance use estimates were higher for most substances among school dropouts, compared to same-aged students. The effects of dropouts on overall estimates increased from the 8th to the 12th grades, as the numbers of dropouts increased. At the 12th grade level, they found that failure to account for dropouts would miss more than half of past year cocaine users, more than half of all lifetime Ecstasy users, 30% of current binge alcohol users, and 25% of current alcohol users.

Because school absenteeism is also known to be associated with increased substance use [22–25], Gfroerer and colleagues [21] additionally investigated the effects of school absenteeism on substance use prevalence estimates in the NSDUH. They reported that those students who missed more days of school were also more likely to be current alcohol users, binge drinkers, and marijuana users. In recognition of this problem, some surveys, such as the YRBS, conduct “make-up” sessions to maximize student opportunities to participate and minimize coverage errors.

3.2. Sampling Errors. Both probability and nonprobability sampling methods are commonly applied in substance use surveys. When probability sampling strategies are employed, all elements within the sample frame have a known, albeit not necessarily equal, probability of selection. The precision of survey statistics derived from such samples can be calculated

with a good degree of confidence and used to estimate the sampling error associated with those statistics. All other things being equal, the size of a random survey sample is inversely associated with the degree of potential sampling error associated with it. The precision of survey estimates also decreases as probability samples deviate from simple random sampling designs, a commonplace occurrence designed to reduce survey costs. Of all the sources of total survey error, the sampling errors related to probability-based sample designs are probably the most well understood, and definable, in practice.

Nonprobability samples are commonly used when research questions focus on special populations believed to be at increased risk for substance use and misuse. There are a variety of well-known nonprobability, or convenience, sample designs commonly used in practice. One of the more popular approaches currently is known as respondent driven sampling (RDS), which was developed by Heckathorn [26, 27] and which has been used in numerous substance use studies [28–30]. Other popular nonprobability strategies in substance use research include venue and facility-based sampling [31–34], snowball sampling [35, 36], time-space sampling [37–39], and advertising for volunteers [40–42]. An important advantage of these designs is their cost effectiveness when researching rare or hidden populations, such as illicit drug users. Because probabilities of selection are unknown, however, there are no definable sampling errors associated with these designs. Rather, nonprobability based sample designs typically suffer from large coverage errors and sampling errors are completely unknown.

3.3. Nonresponse Errors. It is common knowledge that unit response rates in general population surveys have been declining for some time [43–45]. Survey response rates have been historically employed as a proxy indicator of survey quality in general and nonresponse error in particular [46]. Recent research, though, has demonstrated that response rates per se are not necessarily associated with nonresponse bias [47–49]. Rather, it is the degree to which survey respondents and nonrespondents differ from one another in terms of variables of interest to the survey, combined with the survey's response rate that defines nonresponse bias. A British study reported by Plant et al. [50], for example, compared two sets of survey data, with 25% and 79% response rates, respectively. No important differences in self-reports of alcohol consumption were found between the two.

When considering substance use behaviors, there are reasons to be concerned about differences between survey respondents and nonrespondents. Parnanen [4] many years ago suggested that persons who drank heavily might be more difficult to contact as part of survey efforts and would be less likely to cooperate when contacted. In a Canadian survey, De Lint [11] reported that more in-person contact attempts were required to interview those respondents who reported greater numbers of purchases of alcoholic beverages. Cottler et al. [51] additionally reported that those respondents diagnosed with alcohol abuse and dependence required greater numbers of contact attempts in order to complete interviews. Crawford [10] also reported more alcohol consumption among those

respondents most difficult to contact. Using a population register in Sweden, Tibblin [52] found higher rates of survey nonparticipation among middle aged men who were known to have experienced alcohol related problems. There is also some general evidence that survey nonresponse is greater among persons with poor health [53, 54]. A Swedish study has reported that survey respondents were less likely to have been hospitalized with alcohol diagnoses, compared to nonrespondents [55]. These findings are generally interpreted as evidence that heavy drinking may be a barrier to participation in social surveys due to difficulty in making contact and also in convincing those individuals who are contacted to agree to participate [56]. Other investigations, though, have reported no differences in alcohol use between those who do and do not participate in epidemiologic surveys [57–60], and alcohol abstainers have also been found to be underrepresented [56].

It should also be noted that standard field procedures in many surveys actually exclude active substance users from participation. Much research explicitly requires interviewers *not* to conduct interviews with individuals who are visibly intoxicated or appear high on other substances. Kish [61] commented on this problem nearly 50 years ago, referencing a case in which a respondent was drunk by the time they came home after work every day throughout a survey's field period. While such protocols are necessary for orderly data collection and are invoked only infrequently in practice, the potential effects of such protocols on nonresponse bias must nonetheless be considered. In addition, despite some claims to the contrary [62], knowledge that a survey is concerned with substance use appears to have no effect on respondent willingness to participate [55, 63].

Other relevant information comes from studies of attrition in panel surveys, in which the same respondents are interviewed at multiple time points. A number of such investigations have documented higher levels of attrition among high alcohol and drug users [64–73]. In contrast, some other research has found higher attrition among nonusers [74], and Thygesen and colleagues [75] have found both high alcohol intake and abstinence to be associated with increased likelihood of panel attrition. In their study, attrition was also found to be predictive of increased mortality from alcoholic liver cirrhosis and alcoholic liver diseases. Still other research has found no differences between those who do and do not drop out of panel studies [76].

Evidence from research specifically designed to assess nonresponse bias is also informative. Several types of non-response bias studies are routinely conducted. One type is known as follow-up surveys, which typically involve attempting to obtain survey data from nonrespondents to the primary survey [77]. Caspar [57], for example, conducted follow-up face-to-face interviews with a sample of nonrespondents to the 1990 NHSDA, concluding that initial nonrespondents were more likely to report lifetime drug use. Lahaut et al. [56] provide an example of a nonresponse follow-up survey with individuals who initially did not respond to a mail survey and who were subsequently visited by interviewers to complete a face-to-face interview. These analyses suggested that abstainers were underrepresented in

the initial survey. Hill et al. [78] report a telephone follow-up survey of nonrespondents to a primary mail survey. They also found lower reporting of unsafe alcohol consumption among initial nonresponders. Lemmens et al. [60] conducted a telephone follow-up survey of nonrespondents to a face-to-face survey, concluding that there were only small effects of nonresponse on self-reporting of alcohol consumption. An important potential limitation when interpreting findings from follow-up surveys such as these is the use of different modes of data collection between the primary survey and the follow-up effort. Given what is known about mode differences in reporting of substance use behaviors (see Section 4.2), it would not be surprising that a telephone follow-up to a self-administered survey might suggest that the initial survey overestimated substance use, whereas a self-administered nonresponse follow-up survey to an initial interviewer-assisted effort might suggest that it had underestimated substance use. In each case, the effects being attributed to nonresponse bias may actually be a consequence of mode differences rather than systematic nonresponse. Indeed, there are several examples in the literature of surveys that relied on interviewer-assisted follow-up interviews (cf., Hill et al. [78]; Lahaut et al. [56]) that produced data suggesting that primary survey respondents overreport substance use behaviors.

Examples of other types of nonresponse bias analyses that focus on respondent substance use patterns include studies that compare early versus late respondents [56, 79–81]. An example is a study reported by Zhao et al. [62], who compared the answers of persons responding early and late to the Canadian Addictions Survey. Respondents were more likely to have higher incomes and to be educated, males, young adults, and substance users. Such studies employ a continuum of resistance framework that assumes that respondents who require greater effort to contact and interview are more similar to nonrespondents than are those who initially agree to survey requests [82]. Other strategies compare estimates from multiple surveys [62], compare frame data for respondents, nonrespondents, and the full sample [60], or compare estimates from surveys that have high versus low response rates [50].

Another useful strategy for assessing nonresponse bias is to supplement survey data with information obtained from other sources, such as administrative records. For example, Gfroerer et al. [83] examined response patterns in the 1990 National Household Survey on Drug Abuse by merging survey findings with records from the 1990 Decennial Census. Of course, this required special authorization from the government, given the strict data protections associated with the census. They found that persons with some characteristics known to be associated with substance use (i.e., living in urban areas, being male) had lower response rates and that persons with other characteristics believed to be associated with nonsubstance use (older age and higher income levels) also had lower response rates and concluded that these various nonresponse correlates would likely cancel out much of the bias either set might have introduced into the survey estimates.

Finally, it is also important to recognize that high nonresponse rates to individual survey questions (a.k.a., item nonresponse) may also be an indicator of data quality problems in substance use surveys. Some research suggests demographic variability in nonresponse rates to substance use questions. Owens et al. [84] found that African Americans and persons who were separated or divorced were less likely, and females and persons aged 55 and older were more likely, to answer questions concerned with their use illicit drugs. Increased item nonresponse rates to substance use questions among minority groups have also been reported by Witt et al. [85], although Aquilino [86] reported no differences. An item nonresponse study of adolescents additionally found higher nonresponse rates to questions concerned with alcohol and marijuana use among male, compared to female respondents [87].

3.4. Adjustment Errors. Errors of adjustment involve failures to account for the potential effects that a survey's sample design and execution may have upon empirical findings. These may include instances in which sample weights fail to incorporate all sample design and/or nonresponse factors, when variances are unadjusted for the clustering of respondents within sampled geographic areas, or when the available sample weights are not correctly used. An unfortunate example of the failure to properly employ sample weights occurred about a decade ago when a report concerned with illegal sales of alcohol to underage minors in the US seriously overestimated the proportion of all alcohol sales that were reportedly being made to underage youth. The researchers were conducting a secondary analysis of a public release version of the 1998 NHSDA and failed to weigh their data for the survey's stratified sample design, in which young persons aged 12–20 were significantly oversampled. Because only persons under the age of 21 purchase alcohol illegally in the US, their overrepresentation in the unweighted NHSDA data file led to an overrepresentation of illegal sales in those data. This was an error that could have easily been avoided through the use of the preexisting sample weights. The erroneous findings, which were reported nationally, were quickly exposed as flawed [88].

Failure to employ nonresponse weights when survey response rates across different demographic subgroups vary considerably and those same variables which are correlated with substance use patterns can also result in biased substance use estimates. In addition, adjustment errors associated with clustered sample designs (when clustering is not taken into account) can lead to survey estimates with artificially small standard errors that can be misinterpreted as being overly precise [89]. In general, avoidance of adjustment errors would seem to require analysts who possess both substantive knowledge of the addiction processes being examined and methodological knowledge and expertise regarding complex sample design and analysis procedures.

4. Errors of Measurement

4.1. Specification Errors. When survey questionnaires do not correctly conceptualize and/or operationalize constructs of

interest, they are understood to have specification errors. These can take several forms. For example, the street terminology used by drug users is often unique, constantly changing, and varies across locations. Not surprisingly, research demonstrates that the drug names employed in survey questionnaires are not always consistent with the names employed by users in the community [90, 91]. The continued introduction of new substances of course also contributes to specification errors.

In order to adequately assess substance use, it is necessary to ask respondents about all the forms of alcohol and/or drug they may have consumed. Hence, survey questions intended to measure any alcohol or drug use must be able to capture experience with each form of these substances. Global questions that ask about use of the substances in general can be expected to miss some experiences with less common varieties of each. Although these points may seem obvious, they can lead to specification errors more often than most researchers would prefer to admit. Avoiding specification errors requires careful attention during the instrument design process to the specific goals for which the survey is intended to be used.

4.2. Measurement Errors. Measurement error occurs when survey questions fail to measure what they were designed to measure. There are several potential sources of measurement error which must be considered when constructing a survey instrument or analyzing survey data. Broadly speaking, these include design effects, respondent effects, interviewer effects, and context effects.

4.2.1. Design Effects. Virtually every element of a survey that is exposed to respondents is likely to provide them with cues regarding the information being sought [92]. Although many if not most of these cues are unintentional from the researcher's perspective, they can nonetheless be expected to influence self-reports in ways that cannot always be anticipated or controlled. We refer to these as design related errors. Some important design issues discussed below include methods for asking about substance use, mode effects, use of skip patterns, and reference periods. Other design factors that may influence measurement quality include how clearly a survey is introduced as being concerned with substance use, the survey's sponsor, the procedures employed to obtain respondent informed consent, the use of incentives, and the survey's focus as either primarily concerned with substance use or concerned with a more broad set of topics [21]. Regarding this later point, it has been suggested that survey respondents are more willing to discuss negative personal behaviors when they are also asked to report about positive personal behaviors and characteristics [93].

Methods for Asking about Substance Use. Of course, the wording and structure of survey questions can be expected to have a strong influence on the answers obtained, and experimental comparisons have revealed differences in the magnitude of substance use reports obtained using various question measurement strategies. Kroutil et al. [94], for example, have

documented the fact that open-ended questions seriously underestimate drug use prevalence rates. Other research has compared methods for measuring alcohol consumption. Rehm et al. [95] has reported findings from a within-subjects experiment that documents consistently higher prevalence rates for several indicators of harmful drinking when graduated-frequency measures [96] are used, in comparison to the more commonly employed quantity-frequency question response format [97, 98], and weekly drinking recall questions [95]. Other studies have also found graduated-frequency measures to produce higher estimates of alcohol use in comparison to quantity-frequency measures [99, 100]. The superior performance of the graduated-frequency format appears to be based on its ability to more precisely measure irregularly high levels of consumption, although there is some evidence suggesting that the graduated-frequency approach may actually overestimate consumption [100, 101]. Other less commonly used measurement strategies, such as the *yesterday* (or *recent recall*) method of reporting in which respondents are asked to report on their alcohol use during the previous day only, have been found to produce higher estimates than either the quantity-frequency or graduated-frequency measures [102]. The use of a daily diary protocol for collection of alcohol consumption is frequently considered to be a "gold standard" measurement approach [100, 103], but not very practical for most survey applications.

The design of response categories for use in quantity and frequency questions can also influence respondent self-reports. For example, Schwarz [104] has shown how simple changes in the sets of response options presented to respondents, such as emphasizing low versus high frequency events or behaviors, can have effects on overall response patterns. Indeed, Poikolainen and Kärkkäinen [105] have reported obtaining higher alcohol consumption reports when employing quantity and frequency questions that include more heavier intake response options.

It is somewhat ironic that quantity-frequency measures remain commonly utilized in practice, despite the fact that it is conventional wisdom among most substance use researchers that alcohol and drug consumption behaviors are far more variable across even brief time intervals than are assumed by these questions [92, 106]. By their very nature, quantity-frequency items ask for average amounts of use, essentially insuring that they will not capture episodes of heavy or binge drinking. Hasin and Carpenter [107] have documented in a community sample that as many as 30 percent of all respondents report having difficulty when answering typical survey questions concerned with usual drinking patterns due to changes in their drinking behavior during the time period in question and that this problem was particularly acute for persons with symptoms of alcohol dependence. The key advantages of the quantity-frequency measures that make them continue to be popular are their simplicity, ease of answering, and the relatively small amount of space they require in survey instruments. L. C. Sobell and M. B. Sobell [98] and Bloomfield et al. [101] provide comprehensive overviews of the strengths and limitations of various approaches to measuring alcohol consumption in survey questionnaires.

Substance Use Reference Periods. Various reference periods are used to restrict and specify the time intervals for which respondents are asked to retrospectively report their substance use activities. Most often used in practice are 30-day and 12-month reference periods, although there are many variations. Each has its own advantages and disadvantages. It is common knowledge that recall accuracy decays with increasing length of these time intervals [108], as research suggests that greater alcohol prevalence is obtained when shorter reference periods are employed in survey questions [109, 110]. Although more susceptible to recall concerns, a 12-month recall period would have the advantage of being less affected by seasonal variations in substance use [92, 111]. A 30-day reference period, in contrast, might be less likely to capture binge drinking episodes. Hence, some surveys may ask questions about multiple reference periods in order to address the limitations of each.

Also problematic are questions concerned with age of initiation of alcohol and other drug use. Age of first substance use, of course, is considered an important risk factor for subsequent substance abuse, and accurate measurement is hence important [112]. Unfortunately, the length of recall necessary to correctly answer this question can be problematic for many respondents. Forward telescoping, in particular, when respondents underestimate the length of time since an event took place, is an important threat to the quality of self-reports of age of first use [113]. Numerous studies have documented problems with accurate recall of this information [64, 114–121].

Questionnaire Skip Patterns. A common issue when designing substance use questionnaires is the question of whether it is best to employ skip patterns, which allow respondents to avoid answering follow-up questions that are clearly not applicable to them or to instead require all respondents to provide all answers to all items. The rationale for requiring responses to all items is twofold. First, there may be privacy concerns associated with the use of skip patterns, as those who report substance use will require more time to complete all follow-up questions, presumably allowing interviewers and/or other observers to conclude that they are in fact substance users. Second, although it is somewhat burdensome for respondents, it is likely that the presence of skip patterns will be quickly detected by many respondents and possibly motivate some to provide negative answers to filter questions in order to “skip out” longer blocks of questions that request details regarding substance use experiences. As an example of a skip pattern, a question that asks respondents if they have ever used marijuana might be employed as a filter item. Those respondents indicating that they had used marijuana would then be eligible to answer a series of follow-up questions that queried about frequency of use, age of initiation, and so forth. In contrast, avoidance of skip patterns would require respondents to answer all follow-up questions, typically by selecting a “have never used marijuana” response option, which would be available for use with each follow-up question. Such an approach can considerably increase the burden and amount of time necessary to complete a questionnaire for nonusers of the substances being examined. The NSDUH

has historically not employed skip patterns. An experiment reported by Gfroerer et al. [83] investigated the effects of using skip patterns as part of the NHSDA. In their random experiment, they found significantly lower prevalence rates for the five illicit drugs examined when skip patterns were employed. Because no differences were found in alcohol use estimates, it was concluded that privacy concerns associated with answering the most sensitive questions was a more likely explanation for the findings.

Mode Effects. Survey data can be collected using a variety of modalities, including self-administered paper-and-pencil or electronic questionnaires, and telephone or in-person interviews. The presence of mode effects in surveys is well recognized, and there is now a considerable body of evidence documenting the effects of mode on the quality of self-reports of substance use behaviors. In general, survey modes that rely on respondent self-administration are found to obtain greater reports of alcohol and drug use than do those modes that require interviewers to directly ask about use of these substances [55, 58, 122–130]. There is additionally some evidence that these mode effects are greater for more sensitive illicit substances, such as cocaine and marijuana, compared to alcohol use [131].

Among self-administered modes, audio-computer-assisted-self-interviews (ACASI) appear to generate higher reporting of substance use behaviors than do paper-and-pencil (PAPI) self-administered answer sheets [132, 133]. Computer-assisted questionnaires produce data that is more internally consistent and more complete, helping to reduce the need for editing, imputation, and other processing activities that may lead to processing errors (see Section 4.3) [133]. Research has also begun to explore the reliability and validity of substance use surveys conducted via the internet. Eaton and colleagues [134] randomly assigned classes of high school students to respond to PAPI or web questionnaires, concluding that there were few differences in prevalence estimates obtained across the two modes. Ramo and colleagues [135] examined the quality of self-reported marijuana use in a convenience sample of young adults who completed a web-based questionnaire, concluding that such data can be reliably collected. Bauermeister et al. [28] have reported on the use of respondent driven sampling to more systematically sample young adults to participate in a substance use survey. Other investigators have compared internet reporting of alcohol use with reports obtained from self-administered mail questionnaires and both face-to-face and telephone interviews, concluding that online reports have similar levels of measurement quality [136–138].

Among interviewer assisted modes, some evidence suggests that face-to-face interviews appear to produce greater reports than do telephone interviews [86, 123, 139, 140], other evidence suggests no differences in substance use estimates between these two interviewer assisted modes [141, 142], and one study suggests that higher rates of some alcohol-related measures can be obtained by telephone [143]. Some research has also investigated the use of interactive voice recording (IVR) systems (a.k.a., “T-ACASI”—telephone audio

computer-assisted self-interviewing) to improve the quality of substance use data collected by phone [144, 145].

4.2.2. Respondent Effects. Survey respondents vary considerably in their abilities and willingness to provide accurate answers to questions regarding substance use behaviors. Respondent behaviors can be understood within the framework of the generally accepted cognitive model of survey response [146], which recognizes four basic tasks required from respondents when they are answering each survey question. These include (a) question interpretation, (b) memory retrieval, (c) judgment formation, and (d) response editing. This is a useful model for understanding how variability across respondents may influence the quality of self-reported substance use information. Evidence regarding how three of these information processing tasks may influence the quality of substance use behavior reporting is reviewed below.

Question Interpretation. Because respondents sometimes employ substance use terminology that differs from that employed in research questionnaires [91, 147], the risk of miscommunication may be greater in substance use surveys, compared to other topics. The complexity of some substance use terminology may also sometimes lead to respondent confusion. This may be of particular concern in surveys of adolescents, who may not always have sufficient knowledge to correctly respond to questions regarding the use of various drugs [147–149]. Johnston and O'Malley [150] have presented evidence suggesting that respondents sometimes are more likely to deny, or recant, ever having used certain substances that they had previously reported having used (see also additional discussion of recanting in section below on Response Editing). Of particular relevance here is their finding that recanting varies by type of drug being asked about, with the recanting of tranquilizers and barbiturates found to be greater than that for marijuana and cocaine, a finding that they suggest to be related to the complexity of the definitions of these two substances, relative to marijuana and cocaine definitions, which of course also have some complexity. In alcohol research, recent reviews have found that respondents commonly misinterpret standard drink sizes, suggesting that alcohol intake may be systematically underestimated in survey research [151, 152].

A related concern is the degree to which respondent cultural background may influence the interpretation and/or comprehension of survey questions. Substance use patterns and practices are known to vary cross-culturally [153–155], and those varied experiences and beliefs regarding substance use can also be expected to influence respondent knowledge and familiarity with the topic in general and related terminology in particular. Experienced researchers, of course, recognize the importance of investigating and addressing these potential problems by employing focus groups, cognitive interviews, and ethnographic methods during survey development (c.f., Gardner and Tang [156]; Midanik and Hines [157]; Ridolfo [158]; and Thrasher et al. [159]).

Memory Retrieval. The accuracy of respondent recall has been the focus of much attention among methodologists

[160, 161] and has been historically considered one of the more common explanations for inaccurate reporting of substance use behaviors [4, 120, 121]. Indeed, when answering survey questions concerned with substance use, the retrieval of the memories necessary to report accurately can be particularly difficult for several reasons. Poorly worded survey questions may present respondents with difficult cognitive challenges in terms of the effort necessary to retrospectively retrieve specific and/or detailed information that may not be readily accessible in memory [81]. There is also evidence that heavy drinking [4, 162], cocaine [163, 164], and MDMA use [165–167] may be associated with impaired memory. Mensch and Kandel [168] have found inconsistent reporting of marijuana use to be associated with degree of drug use frequency, with the more involved users providing less consistent survey responses, a finding they associate with faulty memory. Although considerable research has been invested in experimenting with strategies for aiding respondents with memory retrieval in general [169, 170], few efforts have focused on aiding recall of substance use information. Hubbard [171], however, has reported a series of experiments that used anchoring manipulations to improve respondent recall, although these were not found to be very effective.

Response Editing. Once respondents have successfully interpreted a survey question and retrieved the relevant information necessary to form an answer, they must decide whether that answer is to be accurately shared with the researcher. Given the illicit and sometimes stigmatizing nature of substance use behaviors, conventional wisdom often suggests that some respondents will make conscious decisions to underreport, or deny altogether, any such behavior [4]. That survey respondents will sometimes attempt to present themselves in a favorable, albeit not completely accurate, light during survey interviews is well understood and is commonly referred to as social desirability bias. Concerns about the potential effects of social desirability bias have been the subject of considerable research in the survey methodology literature [172–175]. In general, respondents are known to overreport socially desirable behaviors, such as voting [176] and exercise [177], and underreport socially undesirable behaviors, including drug and alcohol use [178]. Bradburn and Sudman [172] have explored and documented the sensitive nature of substance use questions by asking a national sample of respondents in the US how uneasy discussing various potentially sensitive topics would make them feel. They found that 42.0 percent reported that they believed most respondents would be “very uneasy” discussing their use of marijuana and that 31.3 and 29.0 percent, respectively, would also be uneasy discussing stimulant and depressant use, and intoxication. Only 10.3 percent indicated that they believed most people would be uneasy discussing drinking in general. This survey, though, was conducted more than 30 years ago and it is unclear to what degree these topics would elicit similar feelings of discomfort today.

Respondents may be uneasy discussing their substance use for several reasons, including the need to avoid the social threat and feelings of shame and embarrassment associated with violating social norms [179, 180]. Reporting

illicit substance use may also be viewed by some respondents as a sign of weakness and, hence, something not to disclose [181]. These points are consistent with research findings that indicate that substance use underreporting increases with the perceived stigma of the substance being discussed [182–184]. Respondents may also elect not to admit to substance use behaviors in order to avoid potential legal sanctions, out of fear that a breach of confidentiality might risk their employment or reputation, and/or because they believe that such information is highly personal and not to be shared. Some research suggests that questions about current use of illicit substances are more likely to produce underestimates when confidentiality is less certain, compared to questions concerned with past use [185]. Experimental studies that have compared substance use reporting patterns when provided with assurance of anonymity versus confidentiality have generally found few differences across conditions [186–188].

Some measures of the propensity to provide socially desirable answers have been found to be associated with substance use reporting such that likelihood of providing socially desirable responses in general is associated with less likelihood of reporting alcohol and/or drug use behavior [172, 189–191]. These findings have been interpreted alternatively as (a) evidence that underreporting of substance use is a consequence of respondent attempts to conceal illicit behavior or as (b) evidence that persons who engage in socially desirable behaviors in general also report, accurately, that they do not engage in substance use behaviors. Although this question remains unresolved, we note that other research has demonstrated the absence of an association between one measure of social desirability, the Crowne-Marlowe scale [173], and a measure of cocaine use underreporting that was based on comparisons of self-reports with biological assays [192].

The accuracy of self-reports of substance use behaviors may also vary by the race/ethnicity of the respondent. A literature review of 36 published studies conducted in the US found consistent evidence of lower reliability and validity rates of substance use reporting among racial and ethnic minority populations [193]. More recent studies have reported similar findings [194, 195]. The specific source of these differences, however, is not clearly understood. Models that have been proposed suggest that greater reporting errors among minority groups may be a consequence of differential group educational achievement and question comprehension, greater minority concerns with privacy, discrimination and risk of prosecution, and/or stronger effects of social desirability pressures on minority groups to report behaviors that conform to majority cultural values. Internationally, cultural differences in normative patterns of alcohol consumption and other substance use may also influence degree of response editing. In nations where wine is considered part of a meal, rather than mood-altering substance, underreporting might be expected to be much less of a concern.

One limitation in much of the research reviewed here is the assumption that greater self-reports of substance use behaviors are more valid [196, 197]. Indeed, overreporting is another measurement concern [197, 198]. There have been cases of respondents providing daily alcohol use reports that

are physically impossible [4]. In surveys of adolescents, there is also a widespread belief that some respondents overreport their alcohol and other drug use, possibly to impress peers and improve one's social status or as part of a general desire for attention [3, 149, 199–202]. Gfroerer and colleagues [21] have speculated that such overreporting of substance use might be more likely to happen during school-based surveys, usually conducted in classroom settings, where peers may be more likely to be aware of respondent answers. It has also been suggested that respondents may in some situations elect to present themselves in a highly negative manner, perhaps for personal amusement or to obtain treatment services [11, 148, 203, 204]. In an effort to identify such overreporters, several investigators have asked respondents about their use of substances that do not exist [205]. It is notable that these studies have found very low self-reported rates of use of these fictitious substances. Petzel et al. [206], for example, found that 4% of his sample of high school students reported the use of the nonexistent drug "bindro." They also found that those who reported the use of a nonexistent drug also reported more use of all other drugs included in their survey, compared to those who indicated, correctly, that they did not use "bindro." Others have reported similar findings when asking survey respondents about the use of nonexistent substances [202, 207–209]. Of course, it may be that heavy drug users just assume, incorrectly, that they have used all available substances at one time or another in their past.

Others have questioned whether or not it is correct to assume that all substance users will be hesitant to accurately report on their patterns of use. Wish et al. [210], for example, have suggested that heavy substance users may be less concerned about social and other consequences of reporting such information. Interviews with persons receiving treatment, though, have found little interest in publicly discussing their patterns of use [211].

Concern with the accuracy of substance use reporting has led to a variety of attempts to validate or corroborate survey responses. For example, several panel surveys have demonstrated considerable stability in respondent reporting of substance use over time [22, 212, 213]. Research, however, has also investigated the recanting of drug and alcohol use, which is the tendency of some panel survey respondents to claim no lifetime experience with a given substance, when they have previously reported having used it [200]. Recanting has been identified in responses to both alcohol [214] and drug use questions [119–121, 150, 201, 215–220]. Depending on the age group being surveyed (adults versus adolescents), recanting may represent deliberate efforts to deny previously reported activity, exaggerations regarding behaviors that never actually took place, poor comprehension of survey questions during at least one wave of interviews, poor recall of information, or simple carelessness when answering [200, 217]. Research by Martino et al. [221] suggests that recanting is a consequence of both deliberate misreporting and errors in understanding of survey questions. In surveys of adolescents, one possible explanation for recanting is that younger and less mature respondents may be more likely to exaggerate substance use during surveys conducted in classroom settings in which peers might be aware of one another's answers

and that they may then provide more accurate answers during subsequent survey waves as they subsequently become more mature [215]. Interestingly, longitudinal follow-ups with Monitoring and the Future Survey respondents have found that recanting is greater among adults with occupations that might be expected to strongly sanction the use of illicit substances, such as those associated with the military and law enforcement [150]. Percy et al. [201] have also documented increased recanting among adolescents who had received drug education during the study period, suggesting a potentially biasing effect of education on self-reports. Higher recanting among low level substance users has also been reported [201, 216].

Other research has sought to validate self-reported substance use behavior by comparing those reports to toxicological findings from biospecimens collected at the time that interviews are conducted. One of the first studies conducted with a community sample (in Chicago) by Fendrich et al. [222] indicated that recent cocaine and heroin use estimates obtained from hair testing were considerably higher than were self-reports obtained from the same respondents. A follow-up survey found that higher rates of cocaine and heroin were obtained from drug assays of hair, saliva, and urine samples, compared to self-reports from respondents to a community survey [178]. A higher estimate of marijuana use, though, was derived from self-reports, compared to drug test assays, a finding that was interpreted as evidence of the limitations of hair testing for the detection of marijuana use. Similar findings of underreporting of cocaine and heroin have also been obtained from general population surveys conducted in Puerto Rico by Colón and colleagues [223, 224] and of men who have sex with men in Chicago [225]. Another study conducted as part of the NSDUH investigated agreement between self-reported use of marijuana and cocaine and urine tests concluded that “most youths aged 12 to 17 and young adults aged 18 to 25 reported their recent drug use accurately” ([226] page 4). Ledgerwood et al. [195] examined the association between hair testing and self-reported illicit drug use, concluding that agreement between tests and self-reports to be substantial for marijuana and cocaine, moderate for opiates, and fair for methamphetamines. Other research has employed urinalysis [227] and hair assays [228] to document drug use frequency underreporting among drug users receiving treatment. While providing valuable insights, it is important to acknowledge that each of these sources of confirmatory biological information is also imperfect measure of substance use, suffering from a variety of limitations, including imprecise and variable detection windows, vulnerability to contamination, and individual and race/ethnic group variability in rates of chemical absorption and retention [229, 230].

Another approach to validating self-reports of substance use is to compare information obtained from respondents with those of significant others, a strategy that has found good but far from perfect levels of corroboration [202, 208, 231–233]. Parents and children have also been asked to corroborate one another’s reports of alcohol use. In a Dutch study, Engels et al. [234] found that both children and parents underestimate one another’s alcohol consumption to

some extent and that underestimation of adolescent alcohol consumption by parents was related to lack of knowledge and control of their children’s activities. An important caveat when employing this approach is that proxy and self-reports generally suffer from the same sources of error [235]. Interestingly, perceptions of untrustworthiness by others have also been found to be associated with drug use recanting among adolescents in a study reported by Weinfurt and Bush [236].

An aggregate level strategy for evaluating self-reports of alcohol use is through comparisons of alcohol sales and tax information. A number of studies have taken this approach and have consistently found evidence suggestive that survey self-reports in some cases vastly underestimate total alcohol consumption [237–240]. State-level estimates from self-reports, though, do correlate fairly strongly with the estimates from sales/tax data, suggesting sensitivity to variations in substance use behavior [238]. One recent study that compared self-reports of alcohol purchases, rather than self-reported alcohol consumption, found much closer agreement between total estimates developed from those self-reports in comparison to total retail alcohol sales in Sweden [241]. Interestingly, this study also found considerable variability by type of alcohol, with sales of wine far more accurately reported than beer and spirits, suggesting the possibility that social desirability concerns may be at least partially responsible, given that wine is likely viewed as a more socially desirable alcoholic beverage, at least in the Swedish context. Reporting of wine consumption was also found to be more complete in a Canadian study [242].

One strategy designed to provide respondents with greater privacy when speaking with interviewers about highly sensitive questions such as substance use behavior is the randomized response technique, first proposed by Warner [243]. Several studies have documented the usefulness of this procedure among both students and adults. Goodstadt and Grusin [244] found higher drug use reporting for five of six substances among high school students in Ontario. Weissman et al. [245] compared substance use self-reports obtained with and without the use of the randomized response technique during telephone interviews conducted as part of a general household survey in New York City and also found increased reporting for three of four substances when using the randomized response technique. An important drawback noted, though, was that only 52% of those randomly assigned to respond using this technique actually agreed to do so. In contrast, McAuliffe et al. [246] reported no differences in reports of illicit drug use among those responding via the randomized response technique, compared to those answering direct questions. Some limitations of this technique include the challenge of correctly administering it in practice and its ability to provide aggregate estimates only [199].

The bogus pipeline is another approach that has been employed in attempts to induce more accurate reporting of substance use behavior. This involves the ethically questionable practice of leading respondents to believe that their questionnaire responses will be validated using some alternative means, when in fact the investigator has no intention of doing so. Rather, the implied threat of validating respondent answers is used to exert pressure on respondents to answer

more truthfully. In general, however, the use of the bogus pipeline procedure has failed to obtain higher estimates of substance use behavior, at least among adolescents [247–249]. A meta-analysis has confirmed the nonefficacy of the bogus pipeline procedure for improved reporting of alcohol consumption and marijuana use [250]. One subsequent study, by Tourangeau et al. [251], did however demonstrate the effectiveness of the bogus pipeline technique for increasing respondent reporting of sensitive behaviors, including alcohol and illicit drug use. In addition, a special population study has suggested that the bogus pipeline procedure may be successful in improving self-reports under certain conditions. Lowe et al. [252] found that, among pregnant women, those randomly assigned to a bogus pipeline condition were nearly twice as likely to report alcohol consumption when completing a self-administered questionnaire.

Finally, when considering respondent related reporting errors, it is highly likely that multiple sources of respondent related reporting errors are operating simultaneously. For example, Johnson and Fendrich [253] demonstrated, using latent measures of cognitive processing difficulties constructed using debriefing probes, that social desirability concerns were predictive of discordant drug use reporting and drug use underreporting, while memory difficulties were predictive of drug use overreporting.

4.2.3. Interviewer Effects. Interviewers can introduce errors by misreading questions, failing to probe answers correctly, not following other elements of standardized survey protocols, and by deliberate falsification of survey interviews [254, 255]. Interviewer affiliation with governmental agencies may also influence respondent willingness to report substance use behaviors [256]. Interestingly and somewhat counterintuitively, interviewers with no prior project-related experience have been found to generate higher levels of marijuana and cocaine reporting in a national substance use survey [130, 257]. Research by Chromy et al. [258] also finds that more experienced interviewers achieve higher response rates, in addition to eliciting fewer reports of substance use behaviors, suggesting they may be more successful in gaining cooperation from nonsubstance users who might find a survey on this topic to be less personally salient or interesting, although they do not believe that this fully accounts for the observed differences, which remain unaccounted for.

Another possible mechanism that may account for interviewer effects involves social distance. It is possible that the social distance between respondents and interviewers may influence respondent willingness to report sensitive behaviors such as substance use. Johnson and colleagues [259] found that adult respondents in a telephone survey regarding substance use treatment needs in Illinois were more likely to report recent and lifetime drug use when respondent-interviewer dyads were characterized as having relatively little social distance. In that study, social distance was measured using a simple count of the number of shared demographic identities (i.e., same gender, same race/ethnicity, similar age, and similar educational attainment). Johnson and colleagues [260] also explored the effects of social distance between race/ethnic groups in a study in

which they probed respondents regarding how comfortable or uncomfortable they would feel when interviewed about their alcohol consumption patterns by interviewers from the same and from other cultural groups. When asked how they would feel if interviewed by an interviewer with the same background, large majorities of African American (88.8%), Mexican American (74.7%), Puerto Rican (85.9%), and non-Hispanic white (92.9%) respondents indicated they would feel comfortable. However, when asked how they would feel if the interviewer asking about their alcohol use was from another cultural group, the proportions indicating they would continue to feel comfortable decreased to 60.0% of African Americans and Mexican Americans and 69.4% of Puerto Ricans. Among non-Hispanic whites, though, the proportion indicating they would continue to be comfortable remained very high (89.3%), suggesting group differences in reactions to interviewers of similar versus different race/ethnic backgrounds.

Other research has also examined the effects on substance use reporting of similarities and differences in various demographic characteristics between interviewers and respondents. In studies conducted in Iowa many years ago, female respondents were more likely to report alcohol consumption to male interviewers, and conversely, male respondents were more likely to report alcohol use to female interviewers [261]. Johnson and Parsons [262] found that homeless respondents were more likely to report drug use to male interviewers, a finding that they linked to a “likely user” hypothesis that suggests that male interviewers were more likely to elicit positive substance use reports because their gender is perceived as being more likely to be substance users themselves and more tolerant of substance use by others. In contrast, a study conducted by Darrow and colleagues [263] reported that gay males were more likely to report drug use to female interviewers, who were viewed as having greater empathy and sympathy for deviant behavior than would male interviewers. In a survey conducted in The Netherlands, higher rates of alcohol use were reported by Turkish and Moroccan respondents to Dutch interviewers, compared to interviewers who were ethnically matched [124]. These researchers also hypothesized that minority respondents may have either (a) exaggerated their alcohol consumption to comply with the perceived norms of the person interviewing them or (b) underreported, or denied altogether, the use of alcohol when interviewed by interviewers from an Islamic background who would have been perceived as having a far less permissive opinion of alcohol use. This limited evidence does not suggest a clear pattern of effects of any interviewer characteristics on respondent self-reports of substance use behaviors, although it does seem likely that interviewer characteristics do matter in many situations.

Interviewer-respondent familiarity with one another may also influence the quality of self-reported substance use behaviors. For example, Mensch and Kandel [168] found that, in the panel survey of National Longitudinal Survey of Youth, marijuana use reporting was lower among respondents who had been interviewed more times previously by the same interviewer, suggesting that interviewer familiarity cued respondents regarding social desirability expectations,

which depressed their drug use reporting. Ironically, again, the use of experienced survey interviewers, something that would typically be considered an important strength of any study, would appear in some circumstances to be a factor contributing to lower quality data, at least when interviewers are serially assigned to the same subsets of respondents.

4.2.4. Context Effects. Various aspects of the social and physical environment within which survey data are collected may also influence the quality of the information collected. One aspect of the social environment that has received attention is the absence or presence of other individuals during the interview, as this is believed to influence the social desirability demands or pressures that respondents may perceive. In general, the presence of others during survey interviews is known to be associated with lower reporting of sensitive behaviors, including substance use. In an early study, Wilson [81] noted that, when interviews were conducted in the presence of another person, average weekly alcohol consumption was lower, compared to interviews conducted in private. Similar findings were reported by Edwards et al. [264], but only among males. Several studies of adolescent reporting of alcohol and drug use also found that the presence of a parent during a household interview reduces respondent willingness to report such behaviors [127, 265–268]. In contrast, Hoyt and Chaloupka [127] also reported that the presence of friends during an interview increased substance use reporting, and Aquilino et al. [266] reported that the presence of a spouse or significant other had no effect on reports of alcohol and drug use. It is important to recognize, though, some potential confounding, as those most likely to have another person present during an interview are those who are married, and those who have children, and these variables are also commonly associated with less substance use behavior.

The physical context within which interviews take place may also influence social desirability pressures and self-report quality. Much of this evidence comes from comparisons of adolescent survey responses when the surveys are completed at home versus in a school setting. In school settings, parental monitoring is likely to be perceived as less of a concern and confidentiality assurance likely to be more credible. Findings support this supposition, as Brener et al. [132] and others [21, 269–271] have reported that adolescents will underreport substance use during household surveys, relative to school-based surveys. Needle and colleagues [272] and Zanes and Matsoukas [273], though, did not find differences in the reports obtained from students in school- versus home-based settings.

4.3. Processing Errors. Once data collection is complete, the construction of a final survey data set requires the implementation of numerous coding and editing rules. The integrity of these rules is particularly critical in substance use surveys, as they typically involve assumptions about the reporting intentions and substance use behaviors of respondents. Fendrich and Johnson [274] have documented important differences in the editing assumptions made across national surveys of

substance use in the US that can substantially influence the prevalence estimates generated by each.

Investigators also use a variety of techniques to screen completed substance use questionnaires for inclusion in final data files. Farrell and colleagues [207] examined the effects of excluding respondents (1) who provided a large number of inconsistent answers and (2) who reported use of a fictitious substance. The effects of excluding these responses on prevalence estimates were considered to be minimal, although they cautioned that exclusionary criteria should be used carefully in order to avoid producing nonrepresentative results.

Also, a past report by the US General Accounting Office [6] identified imputation problems in the National Household Survey on Drug Abuse in which the estimated number of past year heroin users in the US ranged from 232,000 to 701,000, as a consequence of whether missing data imputation procedures were or were not used. The same report also indicated that sample weights used to construct subgroup estimates of the total number of illicit drug users were in some instances based on extremely small numbers of individuals in some weighting cells who reported current drug use. In one case from the 1991 NHSDA, a single 79-year-old woman was projected to represent approximately 142,000 persons believed to have used heroin during the previous year. In such instances, a single erroneous data entry could be expected to have dramatic effects on overall survey estimates.

4.4. Inferential Errors. Inferential errors can be avoided by insuring that the survey questions being employed and the respondents being sampled are representative of the constructs and populations to which the researcher plans to make inferences. To the degree that either the measures or sample fail to represent their intended objects, inferential errors will be realized. Avoiding inferential errors also entails employing sound research designs and appropriate analytic procedures. Experimental findings are considered the strongest evidence for internal validity, and representative samples provide the strongest evidence for external validity. When research designs deviate from these ideals or measures do not adequately assess the constructs of interest, there is a risk of inferential errors that will limit the generalizability of empirical findings. In substance use research, errors of inference can be of several types. Some are a consequence of erroneously concluding that associations between constructs do not exist, due to poor measures and/or research designs. Others involve falsely concluding that associations do exist between constructs when they in fact do not, also as a consequence of inadequate designs and/or measures. The failure to properly adjust a high quality substance use survey for its stratified sample design, discussed earlier in Section 3.4, is an example of an adjustment error that led to a serious inferential error when investigators erroneously concluded that a large fraction of all alcohol sales in the US were being made to underage minors.

5. Discussion

Over several decades, considerable knowledge has been accumulated regarding sources of error in the survey assessment of substance use behaviors. Important gaps remain, however, and continued research will be necessary. Below, I highlight some of the most important questions that I see relevant to each source of survey errors that have been considered in this paper.

Regarding coverage errors, the challenge of constructing representative sample frames for both adolescents and adults continues to increase as electronic communications platforms further diversify. This is a general problem that afflicts all survey research efforts but one that can be particularly problematic for substance use research given the associations between these behaviors and likelihood of being covered by many of the potential sources of sample frames. Identification of supplemental frames that might provide better coverage of heavy substance users and which could be employed, with appropriate weights, as supplements to more traditional sample frames when conducting population surveys should be considered.

When survey estimates are reported, sampling errors, in the form of standard errors or confidence intervals, are commonly included. Although reporting these errors is important to survey transparency, it is important to recognize that sampling errors make strong assumptions that can seldom be met in practice. Most importantly, they assume the absence of all other sources of survey error. Given the unlikelihood of this assumption, merely reporting sampling errors can leave survey consumers with a false sense of the precision of survey estimates, as any sampling errors could be completely overwhelmed by any measurement and/or nonresponse errors, for example, in practice. Understanding how sampling errors in substance use surveys may be influenced by other sources of survey error thus seems to be an important research question to be addressed in the future.

Nonresponse errors seem to be another permanent concern that substance use surveys will need to continually address. Of course, the degree to which nonresponse may bias survey findings will vary from topic to topic and question to question. Given the strong associations detected between substance use and nonresponse patterns, it appears that this error source is also particularly relevant for surveys on this topic. An important issue for additional research is the relative usefulness for substance use surveys of the various nonresponse bias analytic strategies reviewed earlier in this paper. Similarly, research into the relative efficacy of various types of adjustments for nonresponse and other forms of error in substance use surveys would seem to be an important future research topic.

In general, there has been little research into specification errors in substance use surveys. This is an oversight, given general acknowledgment that researchers and potential respondents do not always have a shared understanding of the behaviors being examined. Development of strategies for identifying and investigating potential errors of specification is another research topic in need of attention.

It is my personal opinion that the multiple sources of measurement errors reviewed earlier in this paper pose the greatest threat to the accurate assessment of substance use behaviors. There are several practical questions that remain unresolved, such as the predictive power of social desirability measures, the reasons why experienced interviewers appear to obtain fewer reports of substance use behaviors, and the degree to which adolescents might actually overreport their use of alcohol and/or other drugs. Perhaps even more important, how these widely diverse sets of measurement errors interact with one another is poorly understood and remains largely unexamined. Evaluation of how various sources of measurement errors in substance use surveys interact together to influence survey estimates should be a priority for future research.

In terms of processing errors, surveys concerned with substance use would appear on the surface to be no more vulnerable than other types of survey research. Yet, the complexity of most substance use questionnaires, combined with greater item nonresponse rates in many instances, likely provide greater risks for processing errors that can be linked to complex editing rules and assumptions. A general rule of thumb is that the likelihood of experiencing processing errors is inversely associated with the amount of documentation provided with a survey, as careful documentation is an important indicator of quality research. Continued research into the veracity of data editing decision rules, particularly when handling missing data and/or inconsistent self-reports in substance use surveys, would certainly be welcomed.

As with all other sources of survey related error, inferential errors are not unique to substance use surveys. They are in general a product of poor study design and execution that can seriously limit the value of otherwise commendable efforts. A key to addressing potential inferential errors in research is replication. Study findings take on additional credibility and are accorded stronger inference to the degree that they can be replicated in subsequent investigations. Substance use researchers should seek opportunities to replicate findings from other researchers when conducting their own original studies. And journal editors can provide additional service to science by finding ways to make space available for publishing replication studies that are essential to addressing problems of inferential errors that may otherwise go undetected.

It is important to note that the review presented in this paper was not based on a systematic database search. Rather, it is based on the author's personal familiarity with and experience working with this literature over the past several decades. This should be recognized as a limitation.

Finally, it is strongly recommended that substance use researchers who plan to employ survey research methods recognize and report on their efforts to address each of the potential sources of survey related error discussed in this paper. Developing strategies to systematically and rigorously confront each source of errors and transparently sharing one's successes and failures remains the best approach to minimizing the effects of each when using survey methods to investigate substance use patterns and behaviors.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

References

- [1] M. Hochhauser, "Bias in drug abuse survey research," *International Journal of the Addictions*, vol. 14, no. 5, pp. 675–687, 1979.
- [2] D. Merachnik, "Why initiate a drug survey?" in *Proceedings of the 1st International Conference on Student Drug Surveys*, S. Einstein and S. Allen, Eds., Baywood, Farmingdale, NY, USA, 1972.
- [3] L. Midanik, "The validity of self-reported alcohol consumption and alcohol problems: a literature review," *British Journal of Addiction*, vol. 77, no. 4, pp. 357–382, 1982.
- [4] K. Pernanen, "Validity of survey data on alcohol use," in *Research Advances in Alcohol and Drug Problems*, R. Gibbins, Y. Israel, H. Kalant, R. Popham, W. Schmidt, and R. Smart, Eds., vol. 1, pp. 355–374, John Wiley & Sons, New York, NY, USA, 1974.
- [5] R. E. Popham and W. Schmidt, "Words and deeds: the validity of self-report data on alcohol consumption," *Journal of Studies on Alcohol*, vol. 42, no. 3, pp. 355–358, 1981.
- [6] United States General Accounting Office, "Drug use measurement: strengths, limitations, and recommendations for improvement," Tech. Rep. GAO/PEMD-93-18, Program Evaluation and Methodology Division, Washington, DC, USA, 1993.
- [7] C. Weisner, L. Schmidt, and T. Tam, "Assessing bias in community-based prevalence estimates: towards an unduplicated count of problem drinkers and drug users," *Addiction*, vol. 90, no. 3, pp. 391–405, 1995.
- [8] R. M. Groves, *Survey Errors and Survey Costs*, John Wiley & Sons, New York, NY, USA, 1989.
- [9] P. J. Lavrakas, "Applying a total error perspective for improving research quality in the social, behavioral, and marketing sciences," *Public Opinion Quarterly*, vol. 77, pp. 831–850, 2013.
- [10] A. Crawford, "Bias in a survey of drinking habits," *Alcohol and Alcoholism*, vol. 22, no. 2, pp. 167–179, 1987.
- [11] J. De Lint, "Words and deeds: responses to Popham and Schmidt," *Journal of Studies on Alcohol*, vol. 42, no. 3, pp. 359–361, 1981.
- [12] M. L. Reardon, A. B. Burns, R. Preist, N. Sachs-Ericsson, and A. R. Lang, "Alcohol use and other psychiatric disorders in the formerly homeless and never homeless: prevalence, age of onset, comorbidity, temporal sequencing, and service utilization," *Substance Use & Misuse*, vol. 38, no. 3-6, pp. 601–644, 2003.
- [13] C. D. Delnevo, D. A. Gundersen, and B. T. Hagman, "Declining estimated prevalence of alcohol drinking and smoking among young adults nationally: artifacts of sample undercoverage?" *American Journal of Epidemiology*, vol. 167, no. 1, pp. 15–19, 2008.
- [14] S. J. Blumberg, J. V. Luke, and M. L. Cynamon, "Telephone coverage and health survey estimates: evaluating the need for concern about wireless substitution," *American Journal of Public Health*, vol. 96, no. 5, pp. 926–931, 2006.
- [15] S. J. Blumberg and J. V. Luke, "Reevaluating the need for concern regarding noncoverage bias in landline surveys," *American Journal of Public Health*, vol. 99, no. 10, pp. 1806–1810, 2009.
- [16] M. Livingston, P. Dietze, J. Ferris, D. Pennay, L. Hayes, and S. Lenton, "Surveying alcohol and other drug use through telephone sampling: a comparison of landline and mobile phone samples," *BMC Medical Research Methodology*, vol. 13, no. 1, article 41, 2013.
- [17] S. S. Hu, L. Balluz, M. P. Battaglia, and M. R. Frankel, "Improving public health surveillance using a dual-frame survey of landline and cell phone numbers," *American Journal of Epidemiology*, vol. 173, no. 6, pp. 703–711, 2011.
- [18] S. J. Blumberg and J. V. Luke, *Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July–December 2012*, National Center for Health Statistics, 2013, <http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201306.pdf>.
- [19] E. L. Chavez, R. Edwards, and E. R. Oetting, "Mexican American and white American school dropouts' drug use, health status, and involvement in violence," *Public Health Reports*, vol. 104, no. 6, pp. 594–604, 1989.
- [20] R. C. Swaim, F. Beauvais, E. L. Chavez, and E. R. Oetting, "The effect of school dropout rates on estimates of adolescent substance use among three racial/ethnic groups," *American Journal of Public Health*, vol. 87, no. 1, pp. 51–55, 1997.
- [21] J. Gfroerer, J. Bose, L. Kroutil, M. Lopez, and L. Kann, "Methodological considerations in estimating adolescent substance use," in *Proceedings of the Joint Statistical Meetings, Section on Survey Research Methods*, pp. 4127–4140, 2012.
- [22] J. G. Bachman, L. D. Johnston, and P. M. O'Malley, "Smoking, drinking, and drug use among American high school students: correlates and trends, 1975–1979," *American Journal of Public Health*, vol. 71, no. 1, pp. 59–69, 1981.
- [23] C. D. Cowan, "Coverage, sample design, and weighting in three federal surveys," *Journal of Drug Issues*, vol. 31, no. 3, pp. 599–614, 2001.
- [24] S. Guttmacher, B. C. Weitzman, F. Kapadia, and S. L. Weinberg, "Classroom-based surveys of adolescent risk-taking behaviors: reducing the bias of absenteeism," *American Journal of Public Health*, vol. 92, no. 2, pp. 235–237, 2002.
- [25] D. Kandel, "Reaching the hard-to-reach: illicit drug use among high school absentees," *Addictive Diseases*, vol. 1, no. 4, pp. 465–480, 1975.
- [26] D. D. Heckathorn, "Respondent-driven sampling: a new approach to the study of hidden populations," *Social Problems*, vol. 44, no. 2, pp. 174–199, 1997.
- [27] D. D. Heckathorn, "Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations," *Social Problems*, vol. 49, no. 1, pp. 11–34, 2002.
- [28] J. A. Bauermeister, M. A. Zimmerman, M. M. Johns, P. Glowacki, S. Stoddard, and E. Volz, "Innovative recruitment using online networks: lessons learned from an online study of alcohol and other drug use utilizing a web-based, respondent-driven sampling (webRDS) strategy," *Journal of Studies on Alcohol and Drugs*, vol. 73, no. 5, pp. 834–838, 2012.
- [29] C. McKnight, D. des Jarlais, H. Bramson et al., "Respondent-driven sampling in a study of drug users in New York City: notes from the field," *Journal of Urban Health*, vol. 83, supplement 1, pp. 54–59, 2006.
- [30] J. Ramirez-Valles, L. M. Kuhns, R. T. Campbell, and R. M. Diaz, "Social integration and health: Community involvement, stigmatized identities, and sexual risk in latino sexual minorities," *Journal of Health and Social Behavior*, vol. 51, no. 1, pp. 30–47, 2010.
- [31] M. C. Clatts, L. A. Goldsamt, and H. Yi, "Club drug use among young men who have sex with men in NYC: a preliminary epidemiological profile," *Substance Use and Misuse*, vol. 40, no. 9-10, pp. 1317–1330, 2005.

- [32] P. N. Halkitis, B. N. Fischgrund, and J. T. Parsons, "Explanations for methamphetamine use among gay and bisexual men in New York City," *Substance Use & Misuse*, vol. 40, no. 9-10, pp. 1331-1345, 2005.
- [33] E. Kassira, A. Swetz, R. Bauserman, N. Tomoyasu, E. Caldeira, and L. Solomon, "HIV and AIDS surveillance among inmates in Maryland prisons," *Journal of Urban Health*, vol. 78, no. 2, pp. 256-263, 2001.
- [34] I. Safika, T. P. Johnson, and J. A. Levy, "A venue analysis of predictors of alcohol use prior to sexual intercourse among female sex workers in Senggigi, Indonesia," *International Journal of Drug Policy*, vol. 22, no. 1, pp. 49-55, 2011.
- [35] C. D. Kaplan, D. Korf, and C. Sterk, "Temporal and social contexts of heroin-using populations. An illustration of the snowball sampling technique," *The Journal of Nervous and Mental Disease*, vol. 175, no. 9, pp. 566-574, 1987.
- [36] A. K. Sharma, O. P. Aggarwal, and K. K. Dubey, "Sexual behavior of drug-users: is it different?" *Preventive Medicine*, vol. 34, no. 5, pp. 512-515, 2002.
- [37] M. I. Fernández, G. S. Bowen, L. M. Varga et al., "High rates of club drug use and risky sexual practices among hispanic men who have sex with men in Miami, Florida," *Substance Use and Misuse*, vol. 40, no. 9-10, pp. 1347-1581, 2005.
- [38] F. B. Muhib, L. S. Lin, A. Stueve et al., "A venue-based method for sampling hard-to-reach populations," *Public Health Reports*, vol. 116, no. 1, pp. 216-222, 2001.
- [39] L. A. Valleroy, D. A. Mackellar, J. M. Karon et al., "HIV prevalence and associated risks in young men who have sex with men," *Journal of the American Medical Association*, vol. 284, no. 2, pp. 198-204, 2000.
- [40] S. P. Barrett, S. R. Gross, I. Garand, and R. O. Pihl, "Patterns of simultaneous polysubstance use in Canadian rave attendees," *Substance Use & Misuse*, vol. 40, no. 9-10, pp. 1525-1537, 2005.
- [41] K. B. Levy, K. E. O'Grady, E. D. Wish, and A. M. Arria, "An in-depth qualitative examination of the ecstasy experience: results of a focus group with ecstasy-using college students," *Substance Use and Misuse*, vol. 40, no. 9-10, pp. 1427-1441, 2005.
- [42] K. McElrath, "MDMA and sexual behavior: ecstasy users' perceptions about sexuality and sexual risk," *Substance Use and Misuse*, vol. 40, no. 9-10, pp. 1461-1585, 2005.
- [43] R. M. Groves and M. P. Couper, *Nonresponse in Household Surveys*, John Wiley & Sons, New York, NY, USA, 1998.
- [44] R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little, *Survey Nonresponse*, John Wiley & Sons, New York, NY, USA, 2002.
- [45] C. Steeh, N. Kirgis, B. Cannon, and J. DeWitt, "Are they really as bad as they seem? Nonresponse rates at the end of the 20th century," *Journal of Official Statistics*, vol. 17, pp. 227-247, 2001.
- [46] T. P. Johnson and L. Owens, "Survey response rate reporting in the professional literature," in *Proceedings of the Section on Survey Research Methods*, pp. 127-133, American Statistical Association, Alexandria, Va, USA, 2004.
- [47] R. M. Groves and E. Peytcheva, "The impact of nonresponse rates on nonresponse bias: a meta-analysis," *Public Opinion Quarterly*, vol. 72, no. 2, pp. 167-189, 2008.
- [48] S. Keeter, C. Miller, A. Kohut, R. M. Groves, and S. Presser, "Consequences of reducing nonresponse in a national telephone survey," *Public Opinion Quarterly*, vol. 64, no. 2, pp. 125-148, 2000.
- [49] D. Merkle and M. Edelman, "Nonresponse in exit polls: a comprehensive analysis," in *Survey Nonresponse*, R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little, Eds., pp. 243-258, John Wiley & Sons, New York, NY, USA, 2002.
- [50] M. A. Plant, J. Chick, and N. Kreitman, "The effects of response rates on levels of self-reported alcohol consumption and alcohol-related problems: conclusions from a Scottish study," *British Journal on Alcohol and Alcoholism*, vol. 15, no. 4, pp. 158-163, 1980.
- [51] L. B. Cottler, J. F. Zipp, L. N. Robins, and E. L. Spitznagel, "Difficult-to-recruit respondents and their effect on prevalence estimates in an epidemiologic survey," *American Journal of Epidemiology*, vol. 125, no. 2, pp. 329-339, 1987.
- [52] G. Tibblin, "A population study of 50-year-old men: an analysis of the non-participation group," *Acta Medica Scandinavica*, vol. 178, no. 4, pp. 453-459, 1965.
- [53] G. Cohen and J. C. Duffy, "Are nonrespondents to health surveys less healthy than respondents?" *Journal of Official Statistics*, vol. 18, no. 1, pp. 13-23, 2002.
- [54] N. Hoeymans, E. J. M. Feskens, G. A. M. van Den Bos, and D. Kromhout, "Non-response bias in a study of cardiovascular diseases, functional status and self-rated health among elderly men," *Age and Ageing*, vol. 27, no. 1, pp. 35-40, 1998.
- [55] A. Romelsjö, "The relationship between alcohol consumption and social status in Stockholm. Has the social pattern of alcohol consumption changed?" *International Journal of Epidemiology*, vol. 18, no. 4, pp. 842-851, 1989.
- [56] V. H. M. C. J. Lahaut, H. A. M. Jansen, D. van de Mheen, and H. F. L. Garretsen, "Non-response bias in a sample survey on alcohol consumption," *Alcohol and Alcoholism*, vol. 37, no. 3, pp. 256-260, 2002.
- [57] R. A. Caspar, "Follow-up of non-respondents in 1990," in *Survey Measurement of Drug Use: Methodological Studies*, C. Turner, J. Lessler, and J. Gfroerer, Eds., DHHS Pub. No., ADM 92-1929, pp. 155-173, National Institute on Drug Abuse, Rockville, Md, USA, 1992.
- [58] G. Gmel, "The effect of mode of data collection and of non-response on reported alcohol consumption: a split-sample study in Switzerland," *Addiction*, vol. 95, no. 1, pp. 123-134, 2000.
- [59] L. Iversen and H. Klausen, "Alcohol consumption among laid-off workers before and after closure of a Danish ship-yard: a 2-year follow-up study," *Social Science and Medicine*, vol. 22, no. 1, pp. 107-109, 1986.
- [60] P. H. H. M. Lemmens, E. S. Tan, and R. A. Knibbe, "Bias due to non-response in a Dutch survey on alcohol consumption," *British Journal of Addiction*, vol. 83, no. 9, pp. 1069-1077, 1988.
- [61] L. Kish, *Survey Sampling*, John Wiley & Sons, New York, NY, USA, 1965.
- [62] J. Zhao, T. Stockwell, and S. Macdonald, "Non-response bias in alcohol and drug population surveys," *Drug and Alcohol Review*, vol. 28, no. 6, pp. 648-657, 2009.
- [63] M. A. Plant and T. I. Miller, "Disguised and undisguised questionnaires compared: two alternative approaches to drinking behaviour surveys," *Social Psychiatry*, vol. 12, no. 1, pp. 21-24, 1977.
- [64] S. L. Bailey, R. L. Flewelling, and J. V. Rachal, "The characterization of inconsistencies in self-reports of alcohol and marijuana use in a longitudinal study of adolescents," *Journal of Studies on Alcohol*, vol. 53, no. 6, pp. 636-647, 1992.
- [65] C. M. Beard, A. W. Lane, W. M. O'Fallon, B. L. Riggs, and L. J. Melton III, "Comparison of respondents and nonrespondents in an osteoporosis study," *Annals of Epidemiology*, vol. 4, no. 5, pp. 398-403, 1994.

- [66] K. K. Bucholz, J. J. Shayka, S. L. Marion, C. E. Lewis, E. F. Pribor, and D. M. Rubio, "Is a history of alcohol problems or of psychiatric disorder associated with attrition at 11-year follow-up?" *Annals of Epidemiology*, vol. 6, no. 3, pp. 228–234, 1996.
- [67] R. Caetano, S. Ramisetty-Mikler, and C. McGrath, "Characteristics of non-respondents in a US national longitudinal survey on drinking and intimate partner violence," *Addiction*, vol. 98, no. 6, pp. 791–797, 2003.
- [68] M. Goldberg, J. F. Chastang, M. Zins, I. Niedhammer, and A. Leclerc, "Health problems were the strongest predictors of attrition during follow-up of the GAZEL cohort," *Journal of Clinical Epidemiology*, vol. 59, no. 11, pp. 1213–1221, 2006.
- [69] W. B. Hansen, L. M. Collins, C. K. Malotte, C. A. Johnson, and J. E. Fielding, "Attrition in prevention research," *Journal of Behavioral Medicine*, vol. 8, no. 3, pp. 261–275, 1985.
- [70] T. P. McCoy, E. H. Ip, J. N. Blocker et al., "Attrition bias in a U.S. internet survey of alcohol use among college freshmen," *Journal of Studies on Alcohol and Drugs*, vol. 70, no. 4, pp. 606–614, 2009.
- [71] M. J. Paschall and B. Freisthler, "Does heavy drinking affect academic performance in college? Findings from a prospective study of high achievers," *Journal of Studies on Alcohol*, vol. 64, no. 4, pp. 515–519, 2003.
- [72] D. L. Snow, J. K. Tebes, and M. W. Arthur, "Panel attrition and external validity in adolescent substance use research," *Journal of Consulting and Clinical Psychology*, vol. 60, no. 5, pp. 804–807, 1992.
- [73] T. C. Wild, J. Cunningham, and E. Adlaf, "Nonresponse in a follow-up to a representative telephone survey of adult drinkers," *Journal of Studies on Alcohol*, vol. 62, no. 2, pp. 257–261, 2001.
- [74] M. Garcia, E. Fernandez, A. Schiaffino, C. Borrell, M. Marti, and J. M. Borrás, "Attrition in a population-based cohort eight years after baseline interview: the Cornell Health Interview Survey Follow-up (CHIS.FU) Study," *Annals of Epidemiology*, vol. 15, no. 2, pp. 98–104, 2005.
- [75] L. C. Thygesen, C. Johansen, N. Keiding, E. Giovannucci, and M. Grønbaek, "Effects of sample attrition in a longitudinal study of the association between alcohol intake and all-cause mortality," *Addiction*, vol. 103, no. 7, pp. 1149–1159, 2008.
- [76] B. M. Psaty, A. Cheadle, T. D. Koepsell et al., "Race- and ethnicity-specific characteristics of participants lost to follow-up in a telephone cohort," *American Journal of Epidemiology*, vol. 140, no. 2, pp. 161–171, 1994.
- [77] A. Crawford, "A comparison of participants and non-participants from a British general population survey of alcohol drinking practices," *Journal of the Market Research Society*, vol. 28, pp. 291–297, 1986.
- [78] A. Hill, J. Roberts, P. Ewings, and D. Gunnell, "Non-response bias in a lifestyle survey," *Journal of Public Health Medicine*, vol. 19, no. 2, pp. 203–207, 1997.
- [79] V. M. H. C. J. Lahaut, H. A. M. Jansen, D. van de Mheen, H. F. L. Garretsen, J. E. E. Verdurmen, and A. van Dijk, "Estimating non-response bias in a survey on alcohol consumption: comparison of response waves," *Alcohol & Alcoholism*, vol. 38, no. 2, pp. 128–134, 2003.
- [80] A. M. Trinkoff and C. L. Storr, "Collecting substance use data with an anonymous mailed survey," *Drug and Alcohol Dependence*, vol. 48, no. 1, pp. 1–8, 1997.
- [81] P. Wilson, "Improving the methodology of drinking surveys," *The Statistician*, vol. 30, no. 3, pp. 159–167, 1981.
- [82] I.-F. Lin and N. C. Schaeffer, "Using survey participants to estimate the impact of nonparticipation," *Public Opinion Quarterly*, vol. 59, no. 2, pp. 236–258, 1995.
- [83] J. Gfroerer, J. Lessler, and T. Parsley, "Studies of nonresponse and measurement error in the National Households Survey on Drug Abuse," in *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates*, L. Harrison and A. Hughes, Eds., NIDA Research Monograph 167, NIH Publication No. 97-4147, pp. 273–295, National Institute on Drug Abuse, Rockville, Md, USA, 1997.
- [84] L. Owens, T. P. Johnson, and D. O. Rourke, "Culture and item nonresponse in health surveys," in *Proceedings of the 7th Conference on Health Survey Research Methods*, M. L. Cynamon and R. A. Kulka, Eds., DHHS Publication no. (PHS) 01-1013, pp. 69–74, Hyattsville, Md, USA, 2001.
- [85] M. Witt, J. Pantula, R. Folsom, and C. Cox, "Item nonresponse in 1988," in *Survey Measurement of Drug Use: Methodological Studies*, C. F. Turner, J. T. Lessler, and J. C. Gfroerer, Eds., pp. 85–108, National Institute on Drug Abuse, Rockville, Md, USA, 1988.
- [86] W. S. Aquilino, "Telephone versus face-to-face interviewing for household drug use surveys," *International Journal of the Addictions*, vol. 27, no. 1, pp. 71–91, 1992.
- [87] A. Stueve and L. N. O'Donnell, "Item nonresponse to questions about sex, substance use, and school: results from the Reach for Health Study of African American and Hispanic young adolescents," in *Researching Sexual Behavior: Methodological Issues*, J. H. I. Bancroft, Ed., pp. 376–389, Indiana University Press, Bloomington, Ind, USA, 1997.
- [88] T. Lewin, "Teenage drinking a problem but not in way study found," *New York Times*, 2002.
- [89] S. L. Lohr, *Sampling: Design and Analysis*, Brooks/Cole, Boston, Mass, USA, 2nd edition, 2010.
- [90] M. Hubbard, J. Pantula, and J. Lessler, "Effects of decomposition of complex concepts," in *Survey Measurement of Drug Use*, C. F. Turner, J. T. Lessler, and J. C. Gfroerer, Eds., pp. 245–266, National Institute on Drug Abuse, Rockville, Md, USA, 1992.
- [91] L. J. Ouelett, H. H. Cagle, and D. G. Fisher, "'Crack' versus 'rock' cocaine: the importance of local nomenclature in drug research and education," *Contemporary Drug Problems*, vol. 24, pp. 219–237, 1997.
- [92] T. K. Greenfield and W. C. Kerr, "Alcohol measurement methodology in epidemiology: recent advances and opportunities," *Addiction*, vol. 103, no. 7, pp. 1082–1099, 2008.
- [93] F. J. Fowler and V. L. Stringfellow, "Learning from experience: estimating teen use of alcohol, cigarettes, and marijuana from three survey protocols," *Journal of Drug Issues*, vol. 31, no. 3, pp. 643–664, 2001.
- [94] L. A. Kroutil, M. Vorburger, J. Aldworth, and J. D. Colliver, "Estimated drug use based on direct questioning and open-ended questions: responses in the 2006 National Survey on Drug Use and Health," *International Journal of Methods in Psychiatric Research*, vol. 19, no. 2, pp. 74–87, 2010.
- [95] J. Rehm, T. K. Greenfield, G. Walsh, X. Xie, L. Robson, and E. Single, "Assessment methods for alcohol consumption, prevalence of high risk drinking and harm: a sensitivity analysis," *International Journal of Epidemiology*, vol. 28, no. 2, pp. 219–224, 1999.
- [96] M. E. Hilton, "A comparison of a prospective diary and two summary recall techniques for recording alcohol consumption," *British Journal of Addiction*, vol. 84, no. 9, pp. 1085–1092, 1989.

- [97] D. A. Dawson, "Volume of ethanol consumption: effects of different approaches to measurement," *Journal of Studies on Alcohol*, vol. 59, no. 2, pp. 191–197, 1998.
- [98] L. C. Sobell and M. B. Sobell, "Alcohol consumption measures," in *Assessing Alcohol Problems: A Guide for Clinicians and Researchers*, V. B. Allen and J. P. Wilson, Eds., NIH Pub. No. 03-3745, pp. 75–99, National Institute on Alcohol Abuse and Alcoholism, Bethesda, Md, USA, 2nd edition, 2003.
- [99] L. T. Midanik, "Comparing usual quantity/frequency and graduated frequency scales to assess yearly alcohol consumption: results from the 1990 US national alcohol survey," *Addiction*, vol. 89, no. 4, pp. 407–412, 1994.
- [100] K. Poikolainen, I. Podkletnova, and H. Alho, "Accuracy of quantity-frequency and graduated frequency questionnaires in measuring alcohol intake: Comparison with daily diary and commonly used laboratory markers," *Alcohol & Alcoholism*, vol. 37, no. 6, pp. 573–576, 2002.
- [101] K. Bloomfield, A. Hope, and L. Kraus, "Alcohol survey measures for Europe: a literature review," *Drugs: Education, Prevention and Policy*, vol. 20, no. 5, pp. 348–360, 2013.
- [102] T. Stockwell, J. Zhao, T. Chikritzhs, and T. K. Greenfield, "What did you drink yesterday? Public health relevance of a recent recall method used in the 2004 Australian National Drug Strategy Household Survey," *Addiction*, vol. 103, no. 6, pp. 919–928, 2008.
- [103] K. Poikolainen and P. Kärkkäinen, "Diary gives more accurate information about alcohol consumption than questionnaire," *Drug and Alcohol Dependence*, vol. 11, no. 2, pp. 209–216, 1983.
- [104] N. Schwarz, "Self-reports: how the questions shape the answers," *American Psychologist*, vol. 54, no. 2, pp. 93–105, 1999.
- [105] K. Poikolainen and P. Kärkkäinen, "Nature of questionnaire options affects estimates of alcohol intake," *Journal of Studies on Alcohol*, vol. 46, no. 3, pp. 219–222, 1985.
- [106] L. C. Sobell, T. Cellucci, T. D. Nirenberg, and M. B. Sobell, "Do quantity-frequency data underestimate drinking-related health risks?" *American Journal of Public Health*, vol. 72, no. 8, pp. 823–828, 1982.
- [107] D. Hasin and K. M. Carpenter, "Difficulties with questions on usual drinking and the measurement of alcohol consumption," *Alcoholism: Clinical and Experimental Research*, vol. 22, no. 3, pp. 580–584, 1998.
- [108] R. Tourangeau, "Remembering what happened: memory errors and survey reports," in *The Science of Self Report: Implications for Research and Practice*, A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, and V. S. Cain, Eds., Lawrence Erlbaum Associates, Mahwah, NJ, USA, 2000.
- [109] J. G. Bachman and P. M. O'Malley, "When four months equal a year: inconsistencies in student reports of drug use," *Public Opinion Quarterly*, vol. 45, no. 4, pp. 536–548, 1981.
- [110] J. Simpura and K. Poikolainen, "Accuracy of retrospective measurement of individual alcohol consumption in men; a reinterview after 18 years," *Journal of Studies on Alcohol*, vol. 44, no. 5, pp. 911–917, 1983.
- [111] Y. I. Cho, T. P. Johnson, and M. Fendrich, "Monthly variations in self-reports of alcohol consumption," *Journal of Studies on Alcohol*, vol. 62, no. 2, pp. 268–272, 2001.
- [112] B. F. Grant and D. A. Dawson, "Age of onset of drug use and its association with DSM-IV drug abuse and dependence: results from the national longitudinal alcohol epidemiologic survey," *Journal of Substance Abuse*, vol. 10, no. 2, pp. 163–173, 1998.
- [113] A. M. Shillington, S. I. Woodruff, J. D. Clapp, M. B. Reed, and H. Lemus, "Self-reported age of onset and telescoping for cigarettes, alcohol, and marijuana: across eight years of the National Longitudinal Survey of Youth," *Journal of Child and Adolescent Substance Abuse*, vol. 21, no. 4, pp. 333–348, 2012.
- [114] R. Engels, R. A. Knibbe, and M. J. Drop, "Inconsistencies in adolescents' self-reports of initiation of alcohol and tobacco use," *Addictive Behaviors*, vol. 22, no. 5, pp. 613–623, 1997.
- [115] B. F. Grant, T. C. Harford, D. A. Dawson, P. S. Chou, and R. P. Pickering, "The alcohol use disorder and associated disabilities interview schedule (AUDADIS): reliability of alcohol and drug modules in a general population sample," *Drug and Alcohol Dependence*, vol. 39, no. 1, pp. 37–44, 1995.
- [116] J. A. Humphrey and J. Friedman, "The onset of drinking and intoxication among university students," *Journal of Studies on Alcohol*, vol. 47, no. 6, pp. 455–458, 1986.
- [117] T. P. Johnson and J. A. Mott, "The reliability of self-reported age of onset of tobacco, alcohol and illicit drug use," *Addiction*, vol. 96, no. 8, pp. 1187–1198, 2001.
- [118] J. Prause, D. Dooley, K. A. Ham-Rowbottom, and N. Emptage, "Alcohol drinking onset: a reliability study," *Journal of Child and Adolescent Substance Abuse*, vol. 16, no. 4, pp. 79–90, 2007.
- [119] A. M. Shillington and J. D. Clapp, "Self-report stability of adolescent substance use: are there differences for gender, ethnicity and age?" *Drug and Alcohol Dependence*, vol. 60, no. 1, pp. 19–27, 2000.
- [120] A. M. Shillington, J. D. Clapp, M. B. Reed, and S. I. Woodruff, "Adolescent alcohol use self-report stability: a decade of panel study data," *Journal of Child and Adolescent Substance Abuse*, vol. 20, no. 1, pp. 63–81, 2011.
- [121] A. M. Shillington, M. B. Reed, J. D. Clapp, and S. I. Woodruff, "Testing the length of time theory of recall decay: examining substance use report stability with 10 years of national longitudinal survey of youth data," *Substance Use & Misuse*, vol. 46, no. 9, pp. 1105–1112, 2011.
- [122] W. S. Aquilino and L. A. L. Sciuto, "Effects of interview mode on self-reported drug use," *Public Opinion Quarterly*, vol. 54, no. 3, pp. 362–393, 1990.
- [123] W. S. Aquilino, "Interview mode effects in surveys of drug and alcohol use: a field experiment," *Public Opinion Quarterly*, vol. 58, no. 2, pp. 210–240, 1994.
- [124] A. Dotinga, R. J. J. M. van den Eijnden, W. Bosveld, and H. F. L. Garretsen, "The effect of data collection mode and ethnicity of interviewer on response rates and self-reported alcohol use among Turks and Moroccans in the Netherlands: an experimental study," *Alcohol and Alcoholism*, vol. 40, no. 3, pp. 242–248, 2005.
- [125] J. C. Duffy and J. J. Waterton, "Under-reporting of alcohol consumption in sample surveys: the effect of computer interviewing in fieldwork," *British Journal of Addiction*, vol. 79, no. 3, pp. 303–308, 1984.
- [126] J. Gfroerer and A. Hughes, "Collecting data on illicit drug use by phone," in *Survey Measurement of Drug Use: Methodological Studies*, C. Turner, J. Lessler, and J. Gfroerer, Eds., DHHS Pub. No., ADM 92-1929, pp. 277–295, National Institute on Drug Abuse, Rockville, Md, USA, 1992.
- [127] G. M. Hoyt and F. J. Chaloupka, "Effect of survey conditions on self-reported substance use," *Contemporary Economic Policy*, vol. 12, pp. 109–121, 1994.
- [128] S. Schober, M. F. Caces, M. Pergamit, and L. Branden, "Effects of mode of administration on reporting of drug use in the National

- Longitudinal Survey,” in *Survey Measurement of Drug Use: Methodological Studies*, C. Turner, J. Lessler, and J. Gfroerer, Eds., DHHS Publication No., ADM 92-1929, pp. 267–276, National Institute on Drug Abuse, Rockville, Md, USA, 1992.
- [129] R. Tourangeau and T. W. Smith, “Asking sensitive questions: the impact of data collection mode, question format, and question context,” *Public Opinion Quarterly*, vol. 60, no. 2, pp. 275–304, 1996.
- [130] C. F. Turner, J. T. Lessler, and J. Devore, “Effects of mode of administration and wording on reporting of drug use,” in *Survey Measurement of Drug Use: Methodological Studies*, C. Turner, J. Lessler, and J. Gfroerer, Eds., DHHS Pub. No., ADM 92-1929, pp. 177–220, National Institute on Drug Abuse, Rockville, Md, USA, 1992.
- [131] R. Tourangeau, L. J. Rips, and K. Raskinski, *The Psychology of Survey Response*, Cambridge University Press, Cambridge, UK, 2000.
- [132] N. D. Brener, D. K. Eaton, L. Kann et al., “The association of survey setting and mode with self-reported health risk behaviors among high school students,” *Public Opinion Quarterly*, vol. 70, no. 3, pp. 354–374, 2006.
- [133] J. Chromy, T. Davis, L. Packer, and J. Gfroerer, “Mode effects on substance use measures: comparison of 1999 CAI and PAPI data,” in *Redesigning an Ongoing National Household Survey: Methodological Issues*, J. Gfroerer, J. Eyerman, and J. Chromy, Eds., DHHS Publication No. MA 03-3768, pp. 135–159, Substance Abuse and Mental Health Services Administration, Rockville, Md, USA, 2002.
- [134] D. K. Eaton, N. D. Brener, L. Kann et al., “Comparison of paper-and-pencil versus web administration of the youth risk behavior survey (YRBS): risk behavior prevalence estimates,” *Evaluation Review*, vol. 34, no. 2, pp. 137–153, 2010.
- [135] D. E. Ramo, H. Liu, and J. J. Prochaska, “Reliability and validity of young adults’ anonymous online reports of marijuana use and thoughts about use,” *Psychology of Addictive Behaviors*, vol. 26, no. 4, pp. 801–811, 2012.
- [136] D. A. Hines, E. M. Douglas, and S. Mahmood, “The effects of survey administration on disclosure rates to sensitive items among men: a comparison of an internet panel sample with a RDD telephone sample,” *Computers in Human Behavior*, vol. 26, no. 6, pp. 1327–1335, 2010.
- [137] S. E. McCabe, A. Diez, C. J. Boyd, T. F. Nelson, and E. R. Weitzman, “Comparing web and mail responses in a mixed mode survey in college alcohol use research,” *Addictive Behaviors*, vol. 31, no. 9, pp. 1619–1627, 2006.
- [138] Z. Khadjesari, E. Murray, E. Kalaitzaki et al., “Test-retest reliability of an online measure of past week alcohol consumption (the TOT-AL), and comparison with face-to-face interview,” *Addictive Behaviors*, vol. 34, no. 4, pp. 337–342, 2009.
- [139] J. C. Gfroerer and A. L. Hughes, “The feasibility of collecting drug abuse data by telephone,” *Public Health Reports*, vol. 106, no. 4, pp. 384–393, 1991.
- [140] T. P. Johnson, J. Hougland, and R. Clayton, “Obtaining reports of sensitive behaviors: a comparison of substance use reports from telephone and face-to-face interviews,” *Social Science Quarterly*, vol. 70, pp. 174–183, 1989.
- [141] T. K. Greenfield, L. T. Midanik, and J. D. Rogers, “Effects of telephone versus face-to-face interview modes on reports of alcohol consumption,” *Addiction*, vol. 95, no. 2, pp. 277–284, 2000.
- [142] L. T. Midanik and T. K. Greenfield, “Telephone versus in-person interviews for alcohol use: results of the 2000 National Alcohol Survey,” *Drug and Alcohol Dependence*, vol. 72, no. 3, pp. 209–214, 2003.
- [143] L. T. Midanik, T. K. Greenfield, and J. D. Rogers, “Reports of alcohol-related harm: telephone versus face-to-face interviews,” *Journal of Studies on Alcohol*, vol. 62, no. 1, pp. 74–78, 2001.
- [144] J. N. Gribble, H. G. Miller, P. C. Cooley, J. A. Catania, L. Pollack, and C. F. Turner, “The impact of T-ACASI interviewing on reported drug use among men who have sex with men,” *Substance Use and Misuse*, vol. 35, no. 6–8, pp. 869–1101, 2000.
- [145] M. W. Perrine, J. C. Mundt, J. S. Searles, and L. S. Lester, “Validation of daily self-reported alcohol consumption using interactive voice response (IVR) technology,” *Journal of Studies on Alcohol*, vol. 56, no. 5, pp. 487–490, 1995.
- [146] T. Jabine, M. Straf, J. Tanur, and R. Tourangeau, *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*, National Academy Press, Washington, DC, USA, 1984.
- [147] A. R. Morral, D. F. McCaffrey, and S. Chien, “Measurement of adolescent drug use,” *Journal of Psychoactive Drugs*, vol. 35, no. 3, pp. 301–309, 2003.
- [148] K. M. Harris, B. A. Griffin, D. F. McCaffrey, and A. R. Morral, “Inconsistencies in self-reported drug use by adolescents in substance abuse treatment: implications for outcome and performance measurements,” *Journal of Substance Abuse Treatment*, vol. 34, no. 3, pp. 347–355, 2008.
- [149] H. Swadi, “Validating and improving the validity of self-reports in adolescent substance misuse surveys,” *Journal of Drug Issues*, vol. 20, no. 3, pp. 473–486, 1990.
- [150] L. D. Johnston and P. M. O’Malley, “The recanting or earlier reported drug use by young adults,” in *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates*, L. Harrison and A. Hughes, Eds., NIDA Research Monograph 167, NIH Publication No. 97-4147, pp. 59–80, National Institute on Drug Abuse, Rockville, Md, USA, 1997.
- [151] L. Devos-Comby and J. E. Lange, “‘My drink is larger than yours’? A literature review of self-defined drink sizes and standard drinks,” *Current Drug Abuse Reviews*, vol. 1, no. 2, pp. 162–176, 2008.
- [152] W. C. Kerr and T. Stockwell, “Understanding standard drinks and drinking guidelines,” *Drug and Alcohol Review*, vol. 31, no. 2, pp. 200–205, 2012.
- [153] M. R. de la Rosa and J.-L. R. Adrados, *Drug Abuse Among Minority Youth: Advances in Research and Methodology*, NIDA Research Monograph 130, National Institute on Drug Abuse, Rockville, Md, USA, 1993.
- [154] R. Room, A. Janca, L. A. Bennett et al., “WHO cross-cultural applicability research on diagnosis and assessment of substance use disorders: an overview of methods and selected results,” *Addiction*, vol. 91, no. 2, pp. 199–220, 1996.
- [155] R. Room, “Taking account of cultural and societal influences on substance use diagnoses and criteria,” *Focus*, vol. 5, no. 2, pp. 199–207, 2007.
- [156] B. Gardner and V. Tang, “Reflecting on non-reflective action: an exploratory think-aloud study of self-report habit measures,” *British Journal of Health Psychology*, vol. 19, no. 2, pp. 258–273, 2014.
- [157] L. T. Midanik and A. C. Hines, “‘Unstandard’ ways of answering standard questions: protocol analysis in alcohol survey research,” *Drug and Alcohol Dependence*, vol. 27, no. 3, pp. 245–252, 1991.

- [158] H. Ridolfo, *Testing of the National HIV Behavioral Surveillance System: Results of Interviews Conducted 1/13/2011-4/5/2011*, Questionnaire Design Research Laboratory, National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, Md, USA, 2011, http://www.cdc.gov/qbank/report/Ridolfo_NCHS_2011_NHBSS%20HIV.pdf#page=43.
- [159] J. F. Thrasher, A. C. K. Quah, G. Dominick et al., "Using cognitive interviewing and behavioral coding to determine measurement equivalence across linguistic and cultural groups: an example from the international tobacco control policy evaluation project," *Field Methods*, vol. 23, no. 4, pp. 439–460, 2011.
- [160] W. J. Friedman, "Memory for the time of past events," *Psychological Bulletin*, vol. 113, no. 1, pp. 44–66, 1993.
- [161] S. Sudman, N. M. Bradburn, and N. Schwarz, *Thinking about Answers: The Applications of Cognitive Processes to Survey Methodology*, Jossey-Bass, San Francisco, Calif, USA, 1996.
- [162] T. F. Babor, K. Steinberg, R. Anton, and F. del Boca, "Talk is cheap: measuring drinking outcomes in clinical trials," *Journal of Studies on Alcohol*, vol. 61, no. 1, pp. 55–63, 2000.
- [163] W. G. Van Gorp, J. N. Wilkins, C. H. Hinkin et al., "Declarative and procedural memory functioning in abstinent cocaine abusers," *Archives of General Psychiatry*, vol. 56, no. 1, pp. 85–89, 1999.
- [164] A. Ardila, M. Rosselli, and S. Strumwasser, "Neuropsychological deficits in chronic cocaine abusers," *International Journal of Neuroscience*, vol. 57, no. 1-2, pp. 73–79, 1991.
- [165] K. I. Bolla, U. D. McCann, and G. A. Ricaurte, "Memory impairment in abstinent MDMA ("Ecstasy") users," *Neurology*, vol. 51, no. 6, pp. 1532–1537, 1998.
- [166] M. J. Morgan, "Memory deficits associated with recreational use of "ecstasy" (MDMA)," *Psychopharmacology*, vol. 141, no. 1, pp. 30–36, 1999.
- [167] A. C. Parrott, A. Lees, N. J. Garnham, M. Jones, and K. Wesnes, "Cognitive performance in recreational users of MDMA or "ecstasy": evidence for memory deficits," *Journal of Psychopharmacology*, vol. 12, no. 1, pp. 79–83, 1998.
- [168] B. S. Mensch and D. B. Kandel, "Underreporting of substance use in a national longitudinal youth cohort: individual and interviewer effects," *Public Opinion Quarterly*, vol. 52, no. 1, pp. 100–124, 1988.
- [169] R. F. Belli, *Calendar and Time Diary Methods in Life Course Research*, Sage Publications, Los Angeles, Calif, USA, 2008.
- [170] A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, and V. S. Cain, *The Science of Self-Report: Implications for Research and Practice*, Lawrence Erlbaum Associates, Mahwah, NJ, USA, 2000.
- [171] M. Hubbard, "Laboratory experiments testing new questioning strategies," in *Survey Measurement of Drug Use*, C. F. Turner, J. T. Lessler, and J. C. Gfroerer, Eds., pp. 53–81, National Institute on Drug Abuse, Rockville, Md, USA, 1992.
- [172] N. M. Bradburn and S. Sudman, *Improving Interview Method and Questionnaire Design: Response Effects to Threatening Questions in Survey Research*, Jossey-Bass Publishers, San Francisco, Calif, USA, 1979.
- [173] D. P. Crowne and D. Marlowe, *The Approval Motive: Studies in Evaluative Dependence*, John Wiley & Sons, New York, NY, USA, 1964.
- [174] D. L. Paulhus, "Measurement and control of response bias," in *Measures of Personality and Social Psychological Attitudes*, J. P. Robinson and P. R. Shaver, Eds., pp. 17–59, Academic Press, San Diego, Calif, USA, 1991.
- [175] R. Tourangeau and T. Yan, "Sensitive questions in surveys," *Psychological Bulletin*, vol. 133, no. 5, pp. 859–883, 2007.
- [176] M. W. Traugott and J. P. Katosh, "Response validity in surveys of voting behavior," *Public Opinion Quarterly*, vol. 43, no. 3, pp. 359–377, 1979.
- [177] S. A. Adams, C. E. Matthews, C. B. Ebbeling et al., "The effect of social desirability and social approval on self-reports of physical activity," *American Journal of Epidemiology*, vol. 161, no. 4, pp. 389–398, 2005.
- [178] M. Fendrich, T. P. Johnson, J. S. Wislar, A. Hubbell, and V. Spiehler, "The utility of drug testing in epidemiological research: results from a general population survey," *Addiction*, vol. 99, no. 2, pp. 197–208, 2004.
- [179] M. Harrel, "The validity of self-reported drug use data: the accuracy of responses on confidential self-administered answer sheets," in *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates*, L. Harrison and A. Hughes, Eds., NIDA Research Monograph 167, pp. 37–58, National Institute on Drug Abuse, Rockville, Md, USA, 1997.
- [180] I. Krumpal, "Determinants of social desirability bias in sensitive surveys: a literature review," *Quality and Quantity*, vol. 47, no. 4, pp. 2025–2047, 2013.
- [181] C. Robbins and R. R. Clayton, "Gender-related differences in psychoactive drug use among older adults," *Journal of Drug Issues*, vol. 19, no. 2, pp. 207–219, 1989.
- [182] P. M. O'Malley, J. G. Bachman, and L. D. Johnston, "Reliability and consistency in self-reports of drug use," *The International Journal of the Addictions*, vol. 18, no. 6, pp. 805–824, 1983.
- [183] T. Mieczkowski, "The accuracy of self-reported drug use: an evaluation and analysis of new data," in *Pathways to Criminal Violence*, N. A. Weiner and M. E. Wolfgang, Eds., pp. 275–302, Sage Publications, Newbury Park, Calif, USA, 1989.
- [184] Y. Hser, "Self-reported drug use: results of selected empirical investigations of validity," in *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates*, L. Harrison and A. Hughes, Eds., NIDA Research Monograph 167, NIH Pub No. 97-4147, pp. 320–343, National Institute on Drug Abuse, Rockville, Md, USA, 1997.
- [185] M. J. Luetgert and A. H. Armstrong, "Methodological issues in drug usage surveys: anonymity, recency, and frequency," *International Journal of the Addictions*, vol. 8, no. 4, pp. 683–689, 1973.
- [186] J. H. Malvin and J. M. Moskowitz, "Anonymous versus identifiable self-reports of adolescent drug attitudes, intentions, and use," *Public Opinion Quarterly*, vol. 47, no. 4, pp. 557–566, 1983.
- [187] R. S. Moore and G. M. Ames, "Survey confidentiality vs. anonymity: young men's self-reported substance use," *Journal of Alcohol & Drug Education*, vol. 47, no. 2, pp. 32–41, 2002.
- [188] P. M. O'Malley, L. D. Johnston, J. G. Bachman, and J. Schulenberg, "A comparison of confidential versus anonymous survey procedures: effects on reporting of drug use and related attitudes and beliefs in a national study of students," *Journal of Drug Issues*, vol. 30, no. 1, pp. 35–54, 2000.
- [189] J. H. Pleck, F. L. Sonenstein, and L. Ku, "Black-white differences in adolescent males' substance use: are they explained by underreporting by Blacks?" *Journal of Gender, Culture, and Health*, vol. 1, pp. 247–265, 1996.
- [190] R. G. Watten, "Coping styles in abstainers from alcohol," *Psychopathology*, vol. 29, no. 6, pp. 340–346, 1996.
- [191] J. W. Welte and M. Russell, "Influence of socially desirable responding in a study of stress and substance abuse," *Alcoholism*:

- Clinical and Experimental Research*, vol. 17, no. 4, pp. 758–761, 1993.
- [192] T. P. Johnson, M. Fendrich, and M. E. Mackesy-Amiti, “An evaluation of the validity of the Crowne-Marlowe need for approval scale,” *Quality and Quantity*, vol. 46, no. 6, pp. 1883–1896, 2012.
- [193] T. P. Johnson and P. J. Bowman, “Cross-cultural sources of measurement error in substance use surveys,” *Substance Use and Misuse*, vol. 38, no. 10, pp. 1447–1566, 2003.
- [194] M. Fendrich and T. P. Johnson, “Race/ethnicity differences in the validity of self-reported drug use: Results from a household survey,” *Journal of Urban Health*, vol. 82, no. 3, pp. iii67–iii81, 2005.
- [195] D. M. Ledgerwood, B. A. Goldberger, N. K. Risk, C. E. Lewis, and R. Kato Price, “Comparison between self-report and hair analysis of illicit drug use in a community sample of middle-aged men,” *Addictive Behaviors*, vol. 33, no. 9, pp. 1131–1139, 2008.
- [196] F. K. Del Boca and J. Darkes, “The validity of self-reports of alcohol consumption: state of the science and challenges for research,” *Addiction*, vol. 98, supplement 2, pp. 1–12, 2003.
- [197] P. V. Miller, “Review: is “up” right? The national household survey on drug abuse,” *Public Opinion Quarterly*, vol. 61, no. 4, pp. 627–641, 1997.
- [198] L. D. Harrison, “Understanding the differences in youth drug prevalence rates produced by the MTF, NHSDA, and YRBS studies,” *Journal of Drug Issues*, vol. 31, no. 3, pp. 665–694, 2001.
- [199] N. D. Brener, J. O. G. Billy, and W. R. Grady, “Assessment of factors affecting the validity of self-reported health-risk behavior among adolescents: evidence from the scientific literature,” *Journal of Adolescent Health*, vol. 33, no. 6, pp. 436–457, 2003.
- [200] M. Fendrich, “The undeniable problem of recanting,” *Addiction*, vol. 100, no. 2, pp. 143–144, 2005.
- [201] A. Percy, S. McAlister, K. Higgins, P. McCrystal, and M. Thornton, “Response consistency in young adolescents’ drug use self-reports: a recanting rate analysis,” *Addiction*, vol. 100, no. 2, pp. 189–196, 2005.
- [202] Z. Barnea, G. Rahav, and M. Teichman, “The reliability and consistency of self-reports on substance use in a longitudinal study,” *British Journal of Addiction*, vol. 82, no. 8, pp. 891–898, 1987.
- [203] L. Richter and P. B. Johnson, “Current methods of assessing substance use: a review of strengths, problems, and developments,” *Journal of Drug Issues*, vol. 31, no. 4, pp. 809–832, 2001.
- [204] K. C. Winters, R. D. Stinchfield, G. A. Henly, and R. H. Schwartz, “Validity of adolescent self-report of alcohol and other drug involvement,” *International Journal of the Addictions*, vol. 25, no. 11A, pp. 1379–1395, 1990.
- [205] C. Poulin, P. MacNeil, and W. Mitic, “The validity of a province-wide student drug use survey: lessons in design,” *Canadian Journal of Public Health*, vol. 84, no. 4, pp. 259–264, 1993.
- [206] T. P. Petzel, J. E. Johnson, and J. McKillip, “Response bias in drug surveys,” *Journal of Consulting and Clinical Psychology*, vol. 40, no. 3, pp. 437–439, 1973.
- [207] A. D. Farrell, S. J. Danish, and C. W. Howard, “Evaluation of data screening methods in surveys of adolescents’ drug use,” *Psychological Assessment*, vol. 3, no. 2, pp. 295–298, 1991.
- [208] E. Single, D. Kandel, and B. Johnson, “The reliability and validity of drug use responses in a large scale longitudinal survey,” *Journal of Drug Issues*, vol. 5, pp. 426–443, 1975.
- [209] P. Whitehead and R. Smart, “Validity and reliability of self-reported drug use,” *Canadian Journal of Criminology and Corrections*, vol. 14, pp. 1–8, 1972.
- [210] E. D. Wish, J. A. Hoffman, and S. Nemes, “The validity of self-reports of drug use at treatment admission and at follow-up: comparisons with urinalysis and hair assays,” in *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates*, L. Harrison and A. Hughes, Eds., NIDA Research Monograph 167, NIH Pub No. 97-4147, pp. 200–226, National Institute on Drug Abuse, Rockville, Md, USA, 1997.
- [211] G. Willis, “The use of the psychological laboratory to study sensitive survey topics,” in *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates*, L. Harrison and A. Hughes, Eds., NIDA Research Monograph 167, NIH Pub No. 97-4147, pp. 416–438, National Institute on Drug Abuse, Rockville, Md, USA, 1997.
- [212] A. Merline, J. Jager, and J. E. Schulenberg, “Adolescent risk factors for adult alcohol use and abuse: stability and change of predictive value across early and middle adulthood,” *Addiction*, vol. 103, no. 1, pp. 84–99, 2008.
- [213] D. W. Osgood, L. D. Johnston, P. M. O’Malley, and J. G. Bachman, “The generality of deviance in late adolescence and early adulthood,” *American Sociological Review*, vol. 53, no. 1, pp. 81–93, 1988.
- [214] T. M. Caldwell, B. Rodgers, C. Power, C. Clark, and S. A. Stansfeld, “Drinking histories of self-identified lifetime abstainers and occasional drinkers: findings from the 1958 British Birth Cohort Study,” *Alcohol and Alcoholism*, vol. 41, no. 6, pp. 650–654, 2006.
- [215] M. Fendrich and D. P. Rosenbaum, “Recanting of substance use reports in a longitudinal prevention study,” *Drug and Alcohol Dependence*, vol. 70, no. 3, pp. 241–253, 2003.
- [216] M. Fendrich and C. M. Vaughn, “Diminished lifetime substance use over time: an inquiry into differential underreporting,” *Public Opinion Quarterly*, vol. 58, no. 1, pp. 96–123, 1994.
- [217] M. Fendrich and M. E. Mackesy-Amiti, “Decreased drug reporting in a cross-sectional student drug use survey,” *Journal of Substance Abuse*, vol. 11, no. 2, pp. 161–172, 2000.
- [218] A. M. Shillington, J. D. Clapp, and M. B. Reed, “The stability of self-reported marijuana use across eight years of the National Longitudinal Survey of Youth,” *Journal of Child and Adolescent Substance Abuse*, vol. 20, no. 5, pp. 407–420, 2011.
- [219] A. M. Shillington, S. C. Roesch, M. B. Reed, J. D. Clapp, and S. I. Woodruff, “Typologies of recanting of lifetime cigarette, alcohol and marijuana use during a six-year longitudinal panel study,” *Drug and Alcohol Dependence*, vol. 118, no. 2-3, pp. 134–140, 2011.
- [220] O. Siddiqui, J. A. Mott, T. L. Anderson, and B. R. Flay, “Characteristics of inconsistent respondents who have “ever used” drugs in a school-based sample,” *Substance Use and Misuse*, vol. 34, no. 2, pp. 269–295, 1999.
- [221] S. C. Martino, D. F. McCaffrey, D. J. Klein, and P. L. Ellickson, “Recanting of life-time inhalant use: how big a problem and what to make of it,” *Addiction*, vol. 104, no. 8, pp. 1373–1381, 2009.
- [222] M. Fendrich, T. P. Johnson, S. Sudman, J. S. Wislar, and V. Spiehler, “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*, vol. 149, no. 10, pp. 955–962, 1999.
- [223] H. M. Colón, R. R. Robles, and H. Sahai, “The validity of drug use responses in a household survey in Puerto Rico: comparison of survey responses of cocaine and heroin use with hair tests,” *International Journal of Epidemiology*, vol. 30, no. 5, pp. 1042–1049, 2001.

- [224] H. M. Colón, R. R. Robles, and H. Sahai, "The validity of drug use self-reports among hard core drug users in a household survey in Puerto Rico: comparison of survey responses of cocaine and heroin use with hair tests," *Drug and Alcohol Dependence*, vol. 67, no. 3, pp. 269–279, 2002.
- [225] M. Fendrich, M. E. Mackesy-Amiti, and T. P. Johnson, "Validity of self-reported substance use in men who have sex with men: comparisons with a general population sample," *Annals of Epidemiology*, vol. 18, no. 10, pp. 752–759, 2008.
- [226] L. D. Harrison, S. S. Martin, T. Enev, and D. Harrington, *Comparing Drug Testing and Self-Report of Drug Use among Youths and Young Adults in the General Population*, DHHS Publication No. SMA 07-4249, Methodology Series M-7, Substance Abuse and Mental Health Services Administration, Rockville, Md, USA, 2007.
- [227] A. R. Morral, D. McCaffrey, and M. Y. Iguchi, "Hardcore drug users claim to be occasional users: drug use frequency underreporting," *Drug and Alcohol Dependence*, vol. 57, no. 3, pp. 193–202, 2000.
- [228] K. Tassiopoulos, J. Bernstein, T. Heeren, S. Levenson, R. Hingson, and E. Bernstein, "Predictors of disclosure of continued cocaine use," *Addictive Behaviors*, vol. 31, no. 1, pp. 80–89, 2006.
- [229] K. Wolff, M. Farrell, J. Marsden et al., "A review of biological indicators of illicit drug use, practical considerations and clinical usefulness," *Addiction*, vol. 94, no. 9, pp. 1279–1298, 1999.
- [230] S. F. DeLauder, "Considering issues of racial bias in drug testing where hair is the matrix," *Transforming Anthropology*, vol. 11, no. 2, pp. 54–59, 2004.
- [231] K. Leonard, N. J. Dunn, and T. Jacob, "Drinking problems of alcoholics: correspondence between self and spouse reports," *Addictive Behaviors*, vol. 8, no. 4, pp. 369–373, 1983.
- [232] V. A. Satyanarayana, K. Vaddiparti, P. S. Chandra, C. C. O'Leary, V. Benegal, and L. B. Cottler, "Problem drinking among married men in India: comparison between husband's and wife's reports," *Drug and Alcohol Review*, vol. 29, no. 5, pp. 557–562, 2010.
- [233] L. C. Sobell, S. Agrawal, and M. B. Sobell, "Factors affecting agreement between alcohol abusers' and their collaterals' reports," *Journal of Studies on Alcohol*, vol. 58, no. 4, pp. 405–413, 1997.
- [234] R. C. M. E. Engels, H. Van Der Vorst, M. Deković, and W. Meeus, "Correspondence in collateral and self-reports on alcohol consumption: a within family analysis," *Addictive Behaviors*, vol. 32, no. 5, pp. 1016–1030, 2007.
- [235] F. K. del Boca and J. A. Noll, "Truth or consequences: the validity of self-report data in health services research on addictions," *Addiction*, vol. 95, supplement 3, pp. S347–S360, 2000.
- [236] K. P. Weinfurt and P. J. Bush, "Contradictory subject response in longitudinal research," *Journal of Studies on Alcohol*, vol. 57, no. 3, pp. 273–282, 1996.
- [237] W. C. Kerr and T. K. Greenfield, "Distribution of alcohol consumption and expenditures and the impact of improved measurement on coverage of alcohol sales in the 2000 national alcohol survey," *Alcoholism: Clinical and Experimental Research*, vol. 31, no. 10, pp. 1714–1722, 2007.
- [238] D. E. Nelson, T. S. Naimi, R. D. Brewer, and J. Roeber, "US state alcohol sales compared to survey data, 1993–2006," *Addiction*, vol. 105, no. 9, pp. 1589–1596, 2010.
- [239] J. Rehm, "Measuring quantity, frequency, and volume of drinking," *Alcoholism: Clinical and Experimental Research*, vol. 22, no. 2, pp. 4s–14s, 1998.
- [240] P. F. Smith, P. L. Remington, D. F. Williamson, and R. F. Anda, "A comparison of alcohol sales data with survey data on self-reported alcohol use in 21 states," *American Journal of Public Health*, vol. 80, no. 3, pp. 309–312, 1990.
- [241] M. Ramstedt, "How much alcohol do you buy? A comparison of self-reported alcohol purchases with actual sales," *Addiction*, vol. 105, no. 4, pp. 649–654, 2010.
- [242] M. Sims, "Comparison of sales figures of alcoholic beverages with types and amounts reported by Canadian facts company, limited, in a market survey," Tech. Rep. 1-29-69, Addiction Research Foundation, Toronto, Canada, 1969.
- [243] S. L. Warner, "Randomized response: a survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, vol. 60, no. 309, pp. 63–66, 1965.
- [244] M. S. Goodstadt and V. Grusin, "The randomized response technique: a test on drug use," *Journal of the American Statistical Association*, vol. 70, no. 352, pp. 814–818, 1975.
- [245] A. N. Weissman, R. A. Steer, and D. S. Lipton, "Estimating illicit drug use through telephone interviews and the randomized response technique," *Drug and Alcohol Dependence*, vol. 18, no. 3, pp. 225–233, 1986.
- [246] W. E. McAuliffe, P. Breer, N. W. Ahmadifar, and C. Spino, "Assessment of drug abuser treatment needs in Rhode Island," *American Journal of Public Health*, vol. 81, no. 3, pp. 365–371, 1991.
- [247] P. C. Campanelli, T. E. Dielman, and J. T. Shope, "Validity of adolescents' self-reports of alcohol use and misuse using a bogus pipeline procedure," *Adolescence*, vol. 22, no. 85, pp. 7–22, 1987.
- [248] D. M. Murray and C. L. Perry, "The measurement of substance use among adolescents: when is the 'bogus pipeline' method needed?" *Addictive Behaviors*, vol. 12, no. 3, pp. 225–233, 1987.
- [249] C. E. Werch, D. R. Gorman, P. J. Marty, J. Forbess, and B. Brown, "Effects of the bogus-pipeline on enhancing validity of self-reported adolescent drug use measures," *The Journal of School Health*, vol. 57, no. 6, pp. 232–236, 1987.
- [250] H. Aguinis, C. A. Pierce, and B. M. Quigley, "Enhancing the validity of self-reported alcohol and marijuana consumption using a bogus pipeline procedure: a meta-analytic review," *Basic and Applied Social Psychology*, vol. 16, no. 4, pp. 515–527, 1995.
- [251] T. Tourangeau, T. W. Smith, and K. Rasinski, "Motivation to report sensitive behaviors in surveys: evidence from a bogus pipeline experiment," *Journal of Applied Social Psychology*, vol. 27, pp. 209–222, 1997.
- [252] J. B. Lowe, R. A. Windsor, B. Adams, J. Morris, and Y. Reese, "Use of a bogus pipeline method to increase accuracy of self-reported alcohol consumption among pregnant women," *Journal of Studies on Alcohol*, vol. 47, no. 2, pp. 173–175, 1986.
- [253] T. Johnson and M. Fendrich, "Modeling sources of self-report bias in a survey of drug use epidemiology," *Annals of Epidemiology*, vol. 15, no. 5, pp. 381–389, 2005.
- [254] T. P. Johnson, V. Parker, and C. Clements, "Detection and prevention of data falsification in survey research," *Survey Research*, vol. 32, pp. 1–2, 2001.
- [255] C. F. Turner, J. N. Gribble, A. A. Al-Tayyib, and J. R. Chromy, "Falsification in epidemiologic surveys: detection and remediation [Prepublication Draft]," Technical Papers on Health and Behavior Measurement, No. 53, Research Triangle Institute, Washington, DC, USA, 2002, <http://qcpages.qc.cuny.edu/~cturner/TechPDFs/53.Falsify.pdf>.
- [256] R. A. Grucza, A. M. Abbacchi, T. R. Przybeck, and J. C. Gfroerer, "Discrepancies in estimates of prevalence and correlates of

- substance use and disorders between two national surveys," *Addiction*, vol. 102, no. 4, pp. 623–629, 2007.
- [257] A. Hughes, J. Chrom, K. Giaconetti, and D. Odom, "Impact of interviewer experience on respondent reports of substance use," in *Redesigning an Ongoing National Household Survey: Methodological Issues*, J. Gfroerer, J. Eyerman, and J. Chromy, Eds., DHHS Publication No. SMA 03-3768, pp. 161–184, Substance Abuse and Mental Health Services Administration, Office of Applied Studies, Rockville, Md, USA, 2002.
- [258] J. R. Chromy, J. Eyerman, D. Odom, and A. E. McNeeley, "Association between interviewer experience and substance use prevalence rates in NSDUH," in *Evaluating and Improving Methods Used in the National Survey of Drug Use and Health*, J. Kennet and J. Gfroerer, Eds., DHHS Publication No. SMA 05-4044, Methodology Series M-5, pp. 59–87, Substance Abuse and Mental Health Services Administration, Office of Applied Studies, 2005.
- [259] T. P. Johnson, M. Fendrich, C. Shaligram, A. Garcy, and S. Gillespie, "An evaluation of the effects of interviewer characteristics in an RDD telephone survey of drug use," *Journal of Drug Issues*, vol. 30, no. 1, pp. 77–102, 2000.
- [260] T. Johnson, D. O'Rourke, N. Chavez et al., "Social cognition and responses to survey questions among culturally diverse populations," in *Survey Measurement and Process Quality*, L. Lyberg, P. Biemer, M. Collins et al., Eds., pp. 87–113, John Wiley & Sons, New York, NY, USA, 1997.
- [261] H. A. Mulford and D. E. Miller, "Drinking in Iowa, I. Sociocultural distribution of drinkers," *Quarterly Journal of Studies on Alcohol*, vol. 20, pp. 704–726, 1959.
- [262] T. P. Johnson and J. A. Parsons, "Interviewer effects on self-reported substance use among homeless persons," *Addictive Behaviors*, vol. 19, no. 1, pp. 83–93, 1994.
- [263] W. W. Darrow, H. W. Jaffe, P. A. Thomas et al., "Sex of interviewer, place of interview, and responses of homosexual men to sensitive questions," *Archives of Sexual Behavior*, vol. 15, no. 1, pp. 79–88, 1986.
- [264] S. L. Edwards, M. L. Slattery, and K. N. Ma, "Measurement errors stemming from nonrespondents present at in-person interviews," *Annals of Epidemiology*, vol. 8, no. 4, pp. 272–277, 1998.
- [265] W. S. Aquilino, "Privacy effects on self-reported drug use: interactions with survey mode and respondent characteristics," in *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates*, L. Harrison and A. Hughes, Eds., NIDA Research Monograph 167, pp. 383–415, National Institute on Drug Abuse, Rockville, Md, USA, 1997.
- [266] W. S. Aquilino, D. L. Wright, and A. J. Supple, "Response effects due to bystander presence in CASI and paper-and-pencil surveys of drug use and alcohol use," *Substance Use & Misuse*, vol. 35, no. 6–8, pp. 845–867, 2000.
- [267] J. Gfroerer, "Underreporting of drug use by youths resulting from lack of privacy in household interviews," in *Self-Report Methods of Estimating Drug Use: Meeting Current Challenges to Validity*, B. Rouse, N. Kozel, and L. Richards, Eds., pp. 22–30, National Institute on Drug Abuse, Washington, DC, USA, 1985.
- [268] C. G. Schutz and H. D. Chilcoat, "Breach of privacy in surveys on adolescent drug use: a methodological inquiry," *International Journal of Methods in Psychiatric Research*, vol. 4, pp. 183–188, 1994.
- [269] J. Gfroerer, D. Wright, and A. Kopstein, "Prevalence of youth substance use: the impact of methodological differences between two national surveys," *Drug and Alcohol Dependence*, vol. 47, no. 1, pp. 19–30, 1997.
- [270] L. Kann, N. D. Brener, C. W. Warren, J. L. Collins, and G. A. Giovino, "An assessment of the effect of data collection setting on the prevalence of health risk behaviors among adolescents," *Journal of Adolescent Health*, vol. 31, no. 4, pp. 327–335, 2002.
- [271] I. Rootman and R. G. Smart, "A comparison of alcohol, tobacco and drug use as determined from household and school surveys," *Drug and Alcohol Dependence*, vol. 16, no. 2, pp. 89–94, 1985.
- [272] R. Needle, H. McCubbin, J. Lorence, and M. Hochhauser, "Reliability and validity of adolescent self-reported drug use in a family-based study: a methodological report," *International Journal of the Addictions*, vol. 18, no. 7, pp. 901–912, 1983.
- [273] A. Zanes and E. Matsoukas, "Different settings, different results? A comparison of school and home responses," *Public Opinion Quarterly*, vol. 43, no. 4, pp. 550–557, 1979.
- [274] M. Fendrich and T. P. Johnson, "Examining prevalence differences in three national surveys of youth: impact of consent procedures, mode, and editing rules," *Journal of Drug Issues*, vol. 31, no. 3, pp. 615–642, 2001.