MODELING THE COGNITIVE PROCESSING OF DRUG USE QUESTIONS

Timothy P. Johnson, Survey Research Laboratory and Public Administration Program
Michael Fendrich, Institute for Juvenile Research, Department of Psychiatry
University of Illinois at Chicago

KEY WORDS: Self-Report Validation, Substance Use Surveys, Information Processing

1. INTRODUCTION

Survey measurement error continues to be a problem in the field of substance abuse research. Although epidemiologic surveys retain their status as the primary methodology for monitoring substance use patterns in the U.S., concerns regarding the quality of self-reports of illicit behaviors remain and challenge the credibility of this research (Miller 1997). To date, considerable effort has been invested in assessing potential sources of measurement error and testing innovations designed to improve the quality of self-reports (Harrison and Hughes 1997).

Much of this research has concentrated on evaluating respondent characteristics that may be associated with the accuracy of substance use reporting (Fendrich et al. 1999). In large measure, this work has been driven by the assumption that measurement error in substance use reporting is primarily a result of social desirability concerns. The illegal nature of most recreational drug use and the social stigma often associated with it is believed to provide many respondents with adequate motivation to deliberately underreport, or deny altogether, use of these substances. The desire to maintain a harmonious exchange with an interviewer is viewed as an additional motivation for underreporting. Hence, deliberate underreporting, motivated by confidentiality fears and/or the wish to avoid an uncomfortable social exchange, is widely believed to be the primary mechanism responsible for measurement error in substance abuse research. Schaeffer (2000) has organized a framework for understanding the perceived risks and losses that respondents may associate with answering truthfully when asked about threatening topics such as substance use.

Concerns with privacy, face-saving, and the threat of criminal sanctions also are believed to be the mechanisms underlying apparently robust racial/ethnic differences in the quality of drug use reporting. Johnson and Bowman (2003) recently documented over 30 studies in which the reliability and/or validity of substance use reports varied significantly among survey respondents of differing racial/ethnic backgrounds. In most cases, minority group membership was associated with poorer quality reporting of substance use behaviors. These differences were attributed to a cluster of factors related to social desirability concerns, including greater emphasis on confidentiality, privacy, and harmonious social interactions; greater suspicion of research motives; and greater concern with criminal prosecution in minority communities.

It should be noted that the social desirability framework described above makes the implicit assumption that respondents can accurately comprehend survey questions and retrieve the information necessary to construct correct answers. Indeed, these assumptions appear to be accepted by many in the research community as an article of faith; the relative scarcity of studies that investigate question comprehension and memory retrieval as potential sources of measurement error in substance abuse research seems to support this conclusion. Ironically, the few available studies of these subjects suggest their relevance. Ethnographic work, for example, has documented that the names of drugs communicated in survey questions may not be consistent with the names associated with those drugs in the community (Ouellet, Cagle, and Fisher 1997). This may be a problem somewhat unique to drug abuse research, given ever-changing street drug terminology as new drugs become available and use patterns change. Drug use vocabulary is also likely to vary across regions. Consequently, questionnaire wording may not convey the meaning to respondents that survey researchers assume it does, and personal definitions of various drugs may override those provided in survey questions (Hubbard, Pantula, and Lessler 1992). Other methodological research suggests broad variability in respondent interpretations of survey questions (Schober, Conrad, and Fricker in press; Suessbrick, Schober, and Conrad 2000). Ironically, recent technological innovations designed to improve substance use reporting by reducing social desirability pressures (Turner et al. 1998) themselves may become a barrier to respondent comprehension when interviewers become less available to clarify and help resolve the meanings of objective behavioral questions.

Respondent memory has received considerable attention in recent years as a source of survey measurement error (Sudman, Bradburn, and Schwarz 1996). Although research has been successful in using cognitive interventions to assist respondents in accessing relevant health-related memories and improving responses (Stone et al. 2000), there are few efforts to use knowledge of these processes to improve substance use reporting. Anchoring manipulation experiments reported by Hubbard (1992) achieved mixed success in improving recall.

We are aware of no research that attempts to simultaneously evaluate the effects of these various cognitive processes on substance use reporting error. Doing so would be useful for determining the degree to which the conventional wisdom regarding the primacy of social desirability concerns in substance use reporting is correct and/or the degree to which other elements of information processing, such as question comprehension and memory, contribute to reporting error. The goal of this paper is to examine the relative effects of comprehension, memory, and social desirability on the accuracy of self-reported drug behaviors using a representative community sample. It also will explore potential racial/ethnic differences in these processes.

2. METHODS

The data for this study come from a multistage area probability survey of Chicago residents conducted between June 2001 and January 2002. At stage one, census tracts in Chicago were randomly selected. At stage two, one block was randomly selected from within each sampled tract. At stage three, every household on the sampled block was screened for eligibility. At stage four, one adult age 18–40 was selected at random from within each eligible household (Bryant 1975). Interviews were administered in the home by trained interviewers from the University of Illinois at Chicago Survey Research Laboratory using ACASI procedures. The drug survey portion of the study assessed lifetime and recent drug use using a format similar to that employed by the National Household Survey on Drug Abuse (Substance Abuse and Mental Health Services Administration 2002). Although the overwhelming majority of subjects employed self-administered procedures for the substance use questions (90%), subjects could opt to have their questions administered by the interviewer. All study protocols were reviewed and approved by the University of Illinois at Chicago Institutional Review Board. A total of 627 interviews were completed.

Using American Association of Public Opinion Research (2000) response rate formula #3, the overall response and cooperation rates for the survey were 40% and 59%, respectively. These rates reflect the challenges of conducting in-person survey interviews in urban environments where household response rates tend to be lower for many reasons (Groves and Couper 1998). When restricted access, high-rise apartment buildings were excluded from consideration, the comparable response and cooperation rates were 51% and 80%, respectively. Restricted access apartment buildings are particularly problematic in urban surveys, as a single gatekeeper can successfully decline survey participation on behalf of dozens, and sometimes hundreds, of potential respondents.

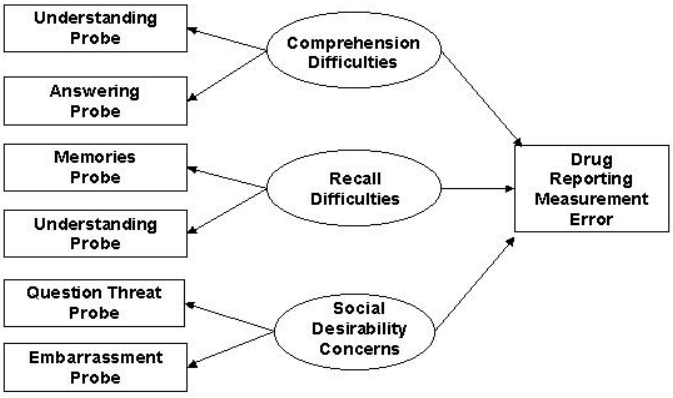
2.1. Questionnaire Contents

In addition to the drug questions noted above, the survey addressed substance use treatment experiences, psychological symptoms, and demographics. Following the main survey, debriefing probes gauging respondent reaction to various aspects of the survey, which are the subject of the present study, were administered (see Table 1). Respondents were asked to answer each of these probes using seven-point Likert-type scales. The use of similar debriefing probes in substance use surveys has been previously reported (Bradburn, Sudman, and Associates 1979; Fendrich, Wislar, and Johnson 2003).

2.2. Drug Testing

Immediately following the drug assessment portion of the survey, subjects were asked to participate in hair, saliva (hereafter referred to as "oral fluid"), and urine testing procedures. Note that subjects were invited to participate in the drug testing after substance use questions had been completed. The strengths and limitations of each of the testing procedures have been discussed elsewhere (Wolf et al. 1999). Specimens were sent to the United States Drug Testing Laboratories in Des Plaines, Illinois, for toxicological analyses and screened for amphetamines, cocaine, marijuana, and opiates. Overall, 91% of the sample participated in at least one drug test. With respect to hair testing, 69% participated, with 12% refusing participation and 21% judged ineligible. Oral fluid test participation was 90%, and urine test participation was 76%.

Figure 1
Conceptual Model of the Relationships Between Information Processing Indicators and Errors in Drug Use Reporting



2.3. Drug Test Classification

Respondents who participated in at least one drug test and were confirmed positive for cocaine, marijuana, or opiates by any of the three tests were classified as drug-test-positive cases. Amphetamine use was dropped from analysis because very few cases tested positive for this substance.

Table 1. Debriefing probes.

	(n)	Mean	SD
Comprehension Probes			
"In general, how much difficulty did you have understanding the drug-related questions included in this survey?"*	(618)	1.71	1.39
"In general, how much difficulty did you have coming up with answers to the drug-related questions in this survey?"*	(619)	1.64	1.21
Memory Probes			
"Please rate how clear your memories were regarding the types of drug-related information asked about in this survey."**	(620)	1.89	1.36
"How certain are you of the accuracy of your answers to these questions?"**	(620)	1.61	0.98
Social Desirability Probes			
"How threatening did you consider the drug-related questions in this survey to be?"***	(618)	1.61	1.26
"How embarrassing did you consider the drug-related questions in this survey to be?"****	(615)	1.57	1.17

* response options: 1 = no difficulty, 7 = a lot of difficulty.
** 1 = very clear, 7 = not at all clear.
*** 1 = not at all threatening, 7 = very threatening.
**** 1 = not at all embarrassing, 7 = very embarrassing.

2.4. Drug Report-Drug Test Concordance

We classified respondents as concordant drug use reporters if their reported past-year use or nonuse of cocaine, marijuana, and opiates were each consistent with the findings of their drug test assays for those substances. Those providing inconsistent information about use of one or more of these substances were classified as being discordant drug use reporters. Overall, 29.0% of the sample for which drug assay data were available (n=568) were classified as discordant drug use reporters. This served as the indicator of drug reporting measurement error for this paper.

2.5. Analysis

We employed covariance structure modeling for these analyses. Using the debriefing probes presented in Table 1, a measurement model was constructed to represent three latent variables: respondent comprehension difficulties, memory difficulties, and social desirability concerns. It was hypothesized that two of the six debriefing items would be useful in representing each of the cognitive processes. The independent effects of each process on the accuracy of substance use reports were simultaneously estimated. Figure 1 presents the conceptual model to be estimated. Multi-group covariance structure modeling was subsequently conducted to examine similarities and differences in the associations between each process and errors in substance use reporting. We constructed the covariance matrices employed in these analyses using Prelis 2 software, and we employed LISREL 8 (Jöreskog and Sörbom 1993a,b) to conduct all covariance structure modeling.

3. RESULTS

Table 1 presents respondent answers to each debriefing probe. As described earlier, each item was measured on a seven-point scale, and all variables were coded such that higher values represented greater levels of self-reported difficulty in responding to the drug use questions included in the survey. Each item was skewed such that respondents tended to indicate few difficulties with each cognitive task. The full seven-point scale range was nonetheless employed by respondents in answering each of these probes.

Table 2. Covariance structure model of the relationships between information processing indicators and errors in drug use reporting (n=555).

	Coefficient	(SE)
A. Measurement Model		
Comprehension		
1. Difficulty understanding drug questions	1.00[a]	---
2. Difficulty answering drug questions	1.16	(0.14)***
Memory		
3. How clear were memories	1.00[a]	---
4. How certain of accuracy of answers	0.74	(0.03)***
Social Desirability		
5. How threatening were drug questions	1.00[a]	---
6. How embarrassing were drug questions	0.92	(0.03)***
B. Structural Model		
1. Effects of comprehension on reporting error	-0.01	(0.04)
2. Effects of memory on reporting error	0.07	(0.03)*
3. Effects of social desirability on reporting error	0.09	(0.03)**

Model fit statistics: R^2 = 0.02; X^2 = 3.72, df = 10, ns; GFI = 1.00; AGFI = 0.99; RMSEA = 0.0; Critical N = 3460.
[a]Fixed parameter.
*p < .05, **p < .01, ***p < .001.

The model depicted in Figure 1 was estimated using all sampled cases for which complete data were available (555 of 627 = 88.5% of all cases). Table 2 presents the coefficients derived from the measurement and structural components of the model, along with several model fit measures. Overall, the data closely fit the specified model (X^2 = 3.72, df = 10, ns; GFI = 1.00; AGFI = 0.99; RMSEA = 0.0; Critical N = 3460). The debriefing probes associated with each latent construct additionally demonstrated a close fit to the specified measurement model (Table 2, Panel A). The structural model (Panel B) revealed significant associations between two of the three latent variables (memory difficulties and social desirability concerns) and discordant drug use reporting.

Sufficient numbers of African-American (n=233), Latino (n=100), and Caucasian (n=168) respondents with complete data permitted cross-group comparisons of the conceptual model using a multi-group covariance structure model. Table 3 presents results of the multi-group model. As Panel A shows, the measurement model appears to have worked well within each of the three racial/ethnic groups examined. The structural model (Panel B) does reveal cross-group differences in the associations between cognitive indicators and errors in drug use measurement. Specifically, social desirability difficulties have a positive effect on measurement error among African-American respondents only. In contrast, memory difficulties are associated with errors in drug use reporting only among Caucasian respondents. None of the cognitive indicators were found to be associated with reporting error among Latino respondents.

Table 3. Multi-group covariance structure model of the relationships between information processing indicators and errors in drug use reporting.

		AFRICAN AMERICANS		CAUCASIANS		LATINOS	
		Coefficient	(SE)	Coefficient	(SE)	Coefficient	(SE)
A.	**Measurement Model**						
	Comprehension						
	1. Understand	1.00[a]	--	1.00[a]	--	1.00[a]	--
	2. Answer	0.80	(0.04)***	2.84	(0.54)***	0.87	(0.07)***
	Memory						
	1. Memory	1.00[a]	--	1.00[a]	--	1.00[a]	--
	2. Certain	1.44	(0.15)***	0.64	(0.05)***	0.76	(0.06)***
	Social Desirability						
	1. Threat	1.00[a]	--	1.00[a]	--	1.00[a]	--
	2. Embarrassment	0.93	(0.05)***	0.94	(0.06)***	0.86	(0.06)***
B.	**Structure Model**						
	1. Effects of Comprehension	-0.01	(0.03)	0.46	(0.26)	0.02	(0.06)
	2. Effects of Memory	-0.05	(0.09)	0.20	(0.07)**	0.02	(0.06)
	3. Effects of Social Desirability	0.24	(0.05)***	0.12	(0.07)	-0.07	(0.06)
C.	**Group Model Statistics**						
	R^2	0.08		0.11		0.01	
	X^2 (df)	8.42		9.62		28.01	
	GFI	0.99		0.98		0.93	
	(n)	(233)		(168)		(100)	

[a]Fixed parameter.
**p < .01, ***p < .001.
X^2 = 46.05; df = 35, ns; RMSEA = 0.04; Critical N = 621.

Table 4. Assessment of nested models designed to test for racial/ethnic group differences in structural parameters.

	Model X^2	(df)	Difference of X^2	(df)
0. Base model: all structural parameters free across racial/ethnic groups	46.05	(35)	---	---
1. Comprehension effects equated across racial/ethnic groups	49.79	(37)	3.74	(2)
2. Memory effects equated across racial/ethnic groups	51.92	(37)	5.87	(2)
3. Social desirability effects equated across racial/ethnic groups	61.92	(37)	15.87**	(2)
4. All three effects equated across racial/ethnic groups	73.65	(41)	27.60***	(6)

*$p < .05$, **$p < .01$, ***$p < .001$.

A series of nested multi-group models were estimated next to verify whether there were in fact racial/ethnic differences in the associations between the indicators of cognitive difficulties and drug use measurement error. In our unconstrained (base) model, all cognitive processes were estimated separately for each racial/ethnic group. In our constrained (nested) models, identical parameters for each specific cognitive process were estimated across all groups. For each cognitive process, a significant difference in the X^2 Goodness of Fit statistic between the unconstrained and constrained model provides a test of the equivalence of that cognitive process across ethnic groups (Hayduck 1987). Table 4 shows the overall X^2 model fit for the base model (described above and shown in Table 3) and each nested model, and the X^2 difference statistic associated with each comparison. The comparison between the base model and a model constraining comprehension effects to be equal across racial/ethnic groups was not significant ($X^2_{df=2} = 3.74$, ns). The comparison between the base model and a model constraining memory effects to be equal across groups was significant at the .10 level ($X^2_{df=2} = 5.89$, $p = <.10$). This suggests a poorer model fit when the effects of memory problems on measurement error are constrained to be equal across groups. When social desirability concerns were constrained, the model fit was significantly worse ($X^2_{df=2} = 15.87$, $p < .01$). Not surprisingly, a final comparison in which all three cognitive indicators were equated across racial/ethnic groups also produced a significantly worse overall model fit ($X^2_{df=6} = 27.60$, $p < .001$).

4. DISCUSSION

Our findings confirm conventional wisdom regarding substance use reporting, which holds that social desirability concerns are the primary source of measurement error when collecting drug use information via survey research. However, subgroup analyses provide some evidence that memory difficulties are also important, although in this case only among Caucasians, who appear to be less vulnerable to social desirability issues than are minority group members (see below). Another interpretation of these findings should be considered as well. Social desirability concerns may have been more evident in this study because they would likely be applicable to very broad sets of drug use questions, making them easily ascertainable via general probes that inquire about "the drug-related questions included in this survey." Both memory and comprehension difficulties, in contrast, are more likely to be question-specific, making them less likely to be captured via the types of generic debriefing probes we employed. Thus, although it appears clear that social desirability concerns are an important source of measurement error in drug use reporting, additional research is needed to more definitively evaluate the degree to which these other dimensions of information processing contribute to data quality. The use of sets of debriefing probes that are question-specific may be one way to avoid this potential problem.

Also consistent with previous speculation (Johnson and Bowman 2003), we found that social desirability concerns are more likely to affect the quality of drug use reports among African Americans than among respondents from other racial/ethnic groups, at least in the U.S. context. Given historical experiences of oppression, discrimination, exploitation (Massey and Denton 1993), suspicions regarding medical researchers' intentions (Friemuth et al. 2001), and the greater risk of criminal prosecution for drug-related offenses (Stone 1998) experienced by African Americans, greater concerns with privacy, confidentiality, and question threat are a rational reaction. Developing culturally sensitive survey methodologies that acknowledge these concerns presents an important research problem that should be addressed. Doing so is likely to be necessary if the current differences in reporting quality across racial/ethnic groups (Johnson and Bowman 2003) are to be eliminated.

Several limitations apply specifically to the interpretation of the causal models presented in this paper. In particular, the dependent measure (concordance) assumes perfect validity for the criterion measure (the drug test); this is clearly not the case. Indeed, drug tests vary with respect to both sensitivity (the ability to detect recent use) and specificity (the ability to accurately screen out non-users). For example, hair tests are insensitive to recent marijuana use. But these limitations are not applicable to just comparisons for marijuana reporting. A number of inconsistencies that might be classified as "overreporting" (or a lack of self-report specificity) are a direct result of the fact that all drug tests can detect only very recent use of a drug. A subject reporting use of *any* substance within the past year who did not actually use that substance within a three-month period for hair, a three-day period for urine, and a 24–48-hour period for oral fluid would have been erroneously classified as a discordant respondent in this study. First, we note that these types of discordant respondents were relatively rare in comparison to underreporters. Further, when we removed over-reporters from the discordant group (analyses not shown), the results paralleled those presented in this paper. Future analyses will spell out the effect of modifications in the construction of the dependent variable in greater detail.

Another important limitation in the criterion measure is that test participation varied considerably by type of test. The test with the most limited window of detection (oral fluid) had the highest participation rate; hair and urine testing had significantly lower participation rates (Fendrich, Johnson, Wislar, and Hubbell in press). This again speaks to the potential limitations in the validity of the criterion measure used for concordance, since most comparisons employed here are based on the least sensitive test.

There also are concerns that survey respondents may not be able to accurately report, or even be aware of, some of the higher order cognitive processes they routinely employ when answering questions (Nisbett and Wilson 1977).

This paper nonetheless highlights a potentially valuable framework for organizing a set of debriefing measures to examine social cognition in survey research. In a previous report, we presented an alternative approach to this same issue (Fendrich et al. 2003). Please note that these two reports provide complementary evidence in support of the utility of these measures. This work suggests that debriefing measures may provide valuable insight about the quality of substance abuse reporting in epidemiologic surveys.

5. ACKNOWLEDGEMENTS

This research was supported by National Institute on Drug Abuse Grant #R01DA12425. The University of Illinois Survey Research Laboratory (SRL) carried out the data collection for this study. We gratefully acknowledge the project support provided for this study by Christine Orland of SRL.

6. REFERENCES

American Association for Public Opinion Research (2000), *Standard Definitions: Final Dispositions of Case Codes and Outcomes for Surveys,* Ann Arbor, MI: Author.

Bradburn, N. M, Sudman, S., and Associates (1979), *Improving Interview Method and Questionnaire Design,* San Francisco: Jossey-Bass.

Bryant, B. E. (1975), "Respondent Selection in a Time of Changing Household Composition," *Journal of Marketing Research,* 12, 129–135.

Fendrich, M., Johnson, T. P., Sudman, S., Wislar, J. S., and Spiehler, V. (1999), "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,* 149, 955–962.

Fendrich, M., Johnson, T. P., Wislar, J. S., and Hubbell, A. (in press), "Drug Test Feasibility in a General Population Household Survey," *Drug and Alcohol Dependence*.

Fendrich, M., Wislar, J. S., and Johnson, T. P. (2003), "The Utility of Debriefing Questions in a Household Survey on Drug Abuse," *Journal of Drug Issues,* 33, 267–284.

Friemuth, V. S., Quinn, S. C., Thomas, S. B., Cole, G., Zook, E., and Duncan, T. (2001), "African Americans' Views on Research and the Tuskegee Syphilis Study," *Social Science & Medicine*, 52, 797–808.

Groves, R. M., and Couper, M. P. (1998), *Nonresponse in Household Interview Surveys*, New York: John Wiley & Sons.

Harrison, L., and Hughes, A. (1997), *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates,* Rockville, MD: National Institute on Drug Abuse.

Hayduck, L. A. (1987), *Structural Equation Modeling with LISREL: Essentials and Advances,* Baltimore: Johns Hopkins University Press.

Hubbard, M. (1992), "Laboratory Experiments Testing New Questioning Strategies," in *Survey Measurement of Drug Use,* eds. C. F. Turner, J. T. Lessler, and J. C. Gfroerer, Rockville, MD: National Institute on Drug Abuse, pp. 53–81.

Hubbard, M., Pantula, J., and Lessler, J. (1992), "Effects of Decomposition of Complex Concepts," in *Survey Measurement of Drug Use,* eds. C. F. Turner, J. T. Lessler, and J. C. Gfroerer, Rockville, MD: National Institute on Drug Abuse, pp. 245–266.

Johnson, T. P., and Bowman, P. J. (2003), "Cross-Cultural Sources of Measurement Error in Substance Use Surveys," *Substance Use & Misuse,* 38, 1447–1490.

Jöreskog, K. G., and Sörbom, D. (1993a), *PRELIS 2 User's Reference Guide,* Chicago: Scientific Software International.

——— (1993b), *LISREL 8 User's Reference Guide*, Chicago: Scientific Software International.

Massey, D. S., and Denton, N. A. (1993), *American Apartheid: Segregation and the Making of the Underclass*, Cambridge: Harvard University Press.

Miller, P. (1997), "Is 'Up' Right? The National Household Survey on Drug Abuse," *Public Opinion Quarterly*, 61, 627–641.

Nisbett, R. E., and Wilson, T. D. (1977), "Telling More Than We Can Know: Verbal Reports on Mental Processes," *Psychological Review*, 84, 231–259.

Ouelett, L. J, Cagle, H. H., and Fisher, D. G. (1997), "'Crack' Versus 'Rock' Cocaine: The Importance of Local Nomenclature in Drug Research and Education," *Contemporary Drug Problems*, 24, 219–237.

Schaeffer, N. C. (2000), "Asking Questions About Sensitive Topics: A Selective Overview," in *The Science of Self-Report: Implications for Research and Practice*, eds. A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, and V. S. Cain, Mahwah, NJ: Lawrence Erlbaum Associates, pp. 105–121.

Schober, M. F., Conrad, F. G., and Fricker, S. S. (in press), "Misunderstanding Standardized Language in Research Interviews," *Applied Cognitive Psychology*.

Stone, C. (1998), "Race, Crime, and the Administration of Justice," *National Institute of Justice Journal*, April, 26–32.

Stone, A. A, Turkkan, J. S, Bachrach, C. A., Jobe, J. B, Kurtzman, H. S., and Cain, V. S. (2000), *The Science of Self-Report: Implications for Research and Practice*, Mahwah, NJ: Lawrence Erlbaum Associates.

Substance Abuse and Mental Health Services Administration, (2002), *Results from the 2001 National Household Survey on Drug Abuse: Volume I, Summary of National Findings* (Office of Applied Studies, NHSDA Series H-17, DHHS Publication No. SMA 02-3758), Rockville, MD.

Sudman, S., Bradburn, N. M., and Schwarz, N. (1996), *Thinking About Answers: The Applications of Cognitive Processes to Survey Methodology,* San Francisco: Jossey-Bass.

Suessbrick, A., Schober, M. F., and Conrad, F. G. (2000), "Different Respondents Interpret Ordinary Questions Quite Differently," *Proceedings of the American Statistical Association, Section on Survey Research Methods,* Alexandria, VA: American Statistical Association, pp. 907–912.

Turner, C. F., Ku, L., Rogers, S. M., Lindberg, L. D., Pleck, J. H., and Sonenstein, F. L. (1998), "Adolescent Sexual Behavior, Drug Use, and Violence: Increased Reporting with Computer Survey Technology," *Science,* 280, 867–873.

Wolf, K., Farrell, M., Marsden, J., Monteiro, M. G., Ali, R., Welch, S., and Strang, J. (1999), "Review of Biological Indicators of Drug Use, Practical Considerations and Clinical Usefulness," *Addiction,* 94, 1279–1298.