ELSEVIER

Social Science & Medicine 61 (2005) 976–984

SOCIAL SCIENCE & MEDICINE

www.elsevier.com/locate/socscimed

Obtaining sensitive information from a wary population: A comparison of telephone and face-to-face surveys of welfare recipients in the United States

William Alex Pridemore[a,*], Kelly R. Damphousse[b], Rebecca K. Moore[c]

[a]*Department of Criminal Justice, Indiana University, Sycamore Hall 302, Bloomington, IN 47405, USA*
[b]*Department of Sociology, University of Oklahoma, Kaufman Hall, Norman, OK 73019, USA*
[c]*Department of Mental Health and Substance Abuse Services, Evaluation and Data Analysis Section, 1200 NE 13th Street, Oklahoma City, OK 73152, USA*

Available online 2 March 2005

Abstract

Recent studies reveal the benefits of technological developments such as audio computer assisted self-interviewing (A-CASI) in interview methodology, especially for surveys of sensitive behavior and information. However, we believe that the selection of mode of administration depends not only on the technology available and the behavior of interest, but also on the specific population under study. We therefore assess survey mode effects on reported rates of alcohol and drug use among welfare recipients, an especially important group for scholars and public health agencies. The sample consisted of adult recipients of Temporary Assistance to Needy Families (TANF) in Oklahoma, January 2001. Adjusting for demographic characteristics, employment, and education, we employ odds ratios to compare 30-day, 1 year, and lifetime prevalence estimates from telephone and face-to-face surveys. Telephone methodology yields similar or higher estimates for *lifetime* prevalence of alcohol, marijuana, and hard drug use and abuse, though lower estimates of *recent* use. We discuss our findings in relation to underfunded public health agencies that must efficiently assess and respond to local levels of drug abuse and we conclude that mode selection may depend upon the population under study.

Keywords: Survey methodology; Alcohol; Drugs; Welfare; USA

Introduction

Substance abuse is a major social and health problem in the United States. According to the 2001 National Household Survey on Drug Abuse (NHSDA), nearly 13 million Americans are classified as heavy drinkers, and in recent years there have been statistically significant increases in the percentage of persons 12 and older using marijuana, cocaine, illegal pain relievers, and tranquilizers (US Department of Health and Human Services, 2003). While surveys like the NHSDA and Monitoring the Future provide estimates of drug use at the national level, treatment agencies rely on information about local populations to estimate alcohol- and drug-related harm, the number of people who might benefit from intervention strategies, and the resources necessary to respond (Crook & Oei, 1998). Estimating levels of drug and alcohol abuse and assessing the need for treatment

*Corresponding author. Tel.: +1 812 856 2220.
E-mail addresses: wpridemo@indiana.edu (W.A. Pridemore), kdamp@ou.edu (K.R. Damphousse).

0277-9536/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.socscimed.2005.01.006

services, however, requires asking individuals about sensitive, socially undesirable, and sometimes illegal behavior. The response bias this can create may be exacerbated when respondents, because of their dependent status, are especially wary of providing truthful information about these behaviors. Arrestees (National Institute of Justice, 2003) and those who receive public aid, for example, are often at an elevated risk for drug abuse but may be less likely to report out of perceived fear of prosecution or losing benefits.

This study evaluates the utility of employing telephone and face-to-face surveys[1] to identify substance abuse among welfare recipients. The extent to which data are useful for estimating the amount and type of treatment needed in a population depends upon their validity. In turn, validity may depend upon interview mode and assurance of confidentiality or anonymity. The latter is especially important here, since our respondents were explicitly informed they were selected for inclusion in the survey because they receive Temporary Assistance to Needy Families (TANF) welfare benefits. Members of this sample were thus likely cautious from first contact: they receive benefits from the state and are being contacted by someone they may assume to be an agent of the state who proceeds to ask them questions about drug use. Many respondents were probably worried that their responses could influence their receipt of benefits or result in other negative ramifications. Yet this is the situation faced by numerous agencies that require a valid assessment of demand for their services when organizing, planning, and budgeting. While questionnaires may be an imperfect tool, they continue to be one of the primary means of collecting this information and are relied upon by public health researchers and agency administrators. Our findings should thus be of immediate and practical benefit.

Literature review

One common mode of data collection employed to gather sensitive information is the telephone survey, which is a preferred method because of cost, ease of administration, convenience, and enhanced quality control via monitoring of interviewers (Lebkowski, 1988). Along with other problems, however, telephone surveys often suffer from low response rates and undercoverage, which may bias estimates of substance use (Greenfield, Midanik, & Rogers, 2000) since this undercoverage is usually non-random with respect to socioeconomic and other characteristics potentially associated with drug use and abuse. Even if a sample is representative of the population under study, however, response bias resulting from survey mode may threaten validity, and Assael and Keon (1982) have shown that non-sampling error contributes more substantially than random sampling error to total design error.

Some studies, for example, have shown that relative to telephone interviews, face-to-face interviews result in increased reporting of drug use (Nebot et al., 1994) and admission of a drinking problem (Mangione, Hingson, & Barrett, 1982). One possible explanation is that face-to-face interviewers are better able to establish rapport and appear more trustworthy than telephone interviewers. McQueen (1989), on the other hand, showed that telephone surveys are more advantageous than face-to-face surveys when measuring other types of sensitive information such as AIDS-related risky behavior. Selecting the appropriate survey mode is crucial since privacy has an effect on response bias when asking questions about sensitive information or socially undesirable behaviors (Johnson, Hougland, & Clayton, 1989; Jones & Forrest, 1992; Turner, Danella, & Rogers, 1995). Recognizing that an affirmative response in the social context of the interview could result in self-stigmatization (DeMaio, 1984), the perceived potential loss of benefits, or some other repercussion, respondents may not present themselves in a negative light (Goffman, 1959) and thus answer untruthfully. As Newman et al. (2002, p. 297) conclude, "Responding to potentially sensitive questions should not be seen as merely 'providing data,' but rather as an activity with complex motivations…includ[ing] maintaining social respect".

Researchers recognize the importance of confidentiality and anonymity, and there has been rapid evolution over the last decade in the use of technology and diverse modes of survey administration in an effort to assure respondents of privacy and thus improve reporting of sensitive behavior. For example, recent research reveals that the use of paper-and-pencil self-administered questionnaires (SAQs) and computer-assisted self-interviewing (CASI) and audio-CASI (A-CASI) tend to result in higher rates of disclosure of sensitive or illegal behavior such as intimate partner violence (Rhodes, Lauderdale, He, Howes, & Levinson, 2002), illicit drug use (Tourangeau & Smith, 1996; Turner et al., 1998), and other stigmatized or risky behaviors (Perlis, Des Jarlais, Friedman, Arasteh, & Turner, 2004).

It would be a mistake, however, to assume that one survey mode is best for all behaviors and populations under study. Cheaper, quicker, and more convenient telephone surveys, for example, may perform just as well as more costly modes of administration when it comes to socially unacceptable behaviors. For example, studies have shown that sensitive information about substance

[1] As discussed in the method section below, the face-to-face surveys were conducted using computer-assisted interviewer administered personal interviews (CAPI). We use the terms "face-to-face" and CAPI interchangeably throughout this paper.

abuse, alcohol consumption, and AIDS-related and other risky behaviors can be obtained effectively via telephone surveys (Aquilino & Wright, 1996; Greenfield et al., 2000; Herzog & Rodgers, 1988; Nebot et al., 1994), presumably because the perceived distance between interviewer and respondent results in the latter being more willing to report sensitive information than in more intimate face-to-face settings (Czaja, 1987).

Other research shows that respondents with lower levels of education and who are less familiar with computers, both of which are correlated with income, may be more averse to the use of (A-)CASI (Couper & Rowe, 1996). Similarly, while paper-and-pencil SAQs are normally perceived by respondents to afford a higher level of confidentiality, they require a certain level of literacy and often contain potentially confusing contingent questioning and skip patterns. Education and income, and thus likely computer familiarity or ability to navigate complex questionnaires, are important characteristics of our sample of TANF recipients (and similar populations served by similar agencies) for which we must estimate substance abuse.

Recent research also shows that the computer, like the telephone, might be *too* impersonal to gain information on certain topics. In these cases, the face-to-face interview may "serve as a medium for interpersonal connection, motivating respondents to express their true problems" (Newman et al., 2002, p. 294). For example, psychiatric symptoms appear to be reported at a higher rate face-to-face than over the telephone (Henson, Cannell, & Roth, 1978), and Newman et al. (2002) show that need for help or treatment is more likely expressed in face-to-face interviews than with A-CASI. Thus, while A-CASI is potentially superior in many respects, it might not be the best choice in all situations (especially when cost is a consideration). Therefore, it is important to evaluate the effectiveness of other techniques to aid our selection of the appropriate mode depending upon the information required and the population under study.

The main goal of the larger project of which this paper is a part is to develop local and statewide estimates of substance abuse among adults receiving TANF benefits in order to plan the amount and type of treatment necessary. These estimates traditionally are generated via telephone surveys. The key assumption about reliability is that respondents reached by telephone are representative of the population of interest, and the key assumption about validity is that the more efficient telephone surveys yield the same data quality as face-to-face surveys. While several studies have evaluated potential mode effects, only a handful have examined them in light of a potentially suspicious population that, because of their dependent status, might possess a heightened sensitivity to questions about socially undesirable or illegal behavior.

Data and method

Sample selection

The population addressed in this study is the 7901 adult TANF recipients in Oklahoma in January 2001. The Oklahoma Department of Human Services (DHS) supplied the research team with a random sample of 2712 people from this population. Each person in the sample was sent a pre-notification letter requesting his/her participation in the study. The letter included the study dates, interviewing hours, assurances of confidentiality, and gave a toll-free number for persons to call to complete the interview if they did not have a telephone (65 respondents used this toll-free number). All participants who completed the telephone survey received a $15 incentive.

Interviewers were instructed to use the list of 2712 phone numbers until they completed 850 interviews. The calls were placed throughout the afternoon and early evening hours 7 days per week. The interviewers completed 853 computer-assisted telephone interviewing (CATI) surveys after attempting to call 1763 phone numbers. Up to 12 call backs were made to those who did not answer the telephone. Only 1274 of the 1763 numbers were valid (475 numbers were not in service, 7 numbers were for institutions/group quarters, and 7 respondents were not currently living in Oklahoma. Using Council of American Survey Research Organizations procedures (CASRO, 2002), the response rate for the telephone survey was 67% (853/1274).

The respondents from the telephone survey who were not contacted (no phone number available, number not in service, no answer, etc.) were combined into a second sample for face-to-face interviewing. Over a 3 month period, interviewers used this second sample to identify and approach the homes of 368 TANF recipients. Recipient homes were visited 7 days per week during daylight hours. The interview team completed 163 interviews using CAPI. The survey was exactly the same as the one administered by telephone to the other respondents in the study. There was no answer at 196 homes, 5 surveys were not completed because of equipment failure, and 4 respondents refused to participate, We are unable to estimate total eligibles for this survey since it is unknown how many of the 196 homes were vacant or no longer housed a TANF recipient. Of the 172 TANF recipients who *were* at home when the interviewer approached, however, there was a 95% completion rate (172–5 (computer failure)–4 (refused to take part) = 163/172).

We thus have a sample of CATI phone respondents ($n = 853$) and a sample of CAPI face-to-face respondents ($n = 163$). It is important to point out that we did not create two random samples (one for each survey mode) that were similar in size and composition, as

might be expected in a more traditional study. Instead we simulated the very realistic situation faced by countless local agencies across the nation when attempting to meet their mission of assessing and responding to local levels of drug abuse and need for treatment.

Overview of survey and method

Data were collected on several demographic characteristics, including sex, age, race, marital status, education, and work status, in order to compare the two samples and to control for these factors when evaluating mode effects. A series of questions assessed the extent of substance use and abuse among respondents. The substances examined included cigarettes, alcohol, and several illicit drugs. Respondents were asked if they had *ever* smoked cigarettes, if they had done so within the last year, or in the past 30 days. The *amount* of cigarette use was not measured. Respondents were also asked if they had ever consumed *any* alcohol, if they had done so within the last year, or in the past 30 days. If they had consumed alcohol in the past, they were also asked if they had ever consumed more than 4 alcoholic drinks in one sitting (5 drinks for males), if they had done so within the last year, or in the past 30 days. We defined this as "binge drinking". We also asked if the respondents had ever consumed alcohol for 2 days without sobering up, if they had done so within the last year, or in the past 30 days. We defined this as "binging". Finally, we asked each respondent if they felt like they had a "problem" with alcohol. Similar questions were asked about marijuana, cocaine, heroin, sedatives, stimulants, hallucinogens, and intravenous drug use.

As with any study of this type, we assume that higher prevalence is associated with less under-reporting. We dichotomized all the responses into yes/no (e.g., did the respondent report smoking cigarettes, drinking more than five drinks in one sitting, using cocaine, ever injecting any drug to get high?). Following the analytical strategy of prior studies, we used logistic regression to evaluate mode effects on prevalence estimates, first creating crude odds ratios and then estimating adjusted odds ratios to control for the distribution of demographic characteristics in the two samples.

Results

Table 1 presents the distribution of demographic characteristics of respondents in the two samples. Overall, the TANF subjects are largely young, non-white, not employed at least part-time, and overwhelmingly female and unmarried. While generally similar in composition, relative to the telephone sample the face-to-face sample consists of proportionally fewer respondents that are white non-Hispanic, single/divorced, and unemployed. These differences are statistically significant only for marital and employment status.

Table 1
Characteristics of telephone ($n = 853$) and face-to-face ($n = 163$) samples

Characteristic	Proportion of telephone sample	Proportion of face-to-face sample	*p*-Value
Sex			
Female	93.6	90.2	.822
Age			
18–29	48.3	49.1	.177
Race			
White non-Hispanic	46.3	38.7	.075
Marital status			
Never married or divorced/separated	84.5	75.2	.008
Education			
Did not complete high school	24.0	28.3	.261
Work status			
Not currently employed at least part-time	70.2	57.4	.003

Original response categories were more refined than we report here. This table shows only the collapsed categories as they are controlled for the adjusted odds ratios.

Tables 2 and 3 present estimates by mode of administration of lifetime, 30-day, and 1-year prevalence, as well as if the respondent believes s/he might ever have had a substance problem. Table 2 shows results for cigarettes, alcohol, and marijuana. There are no differences between the two samples in cigarette use, which is expected given widespread use and the lack of stigma attached to smoking as a problem behavior relative to the other substances in this survey. This also provides further confidence in the relative similarity of the two samples. Controlling for demographic characteristics, face-to-face respondents were significantly more likely to admit using alcohol in the last 30 days (OR = 1.48, $p = .03$), ever drinking at least 2 days straight without sobering up (OR = 1.60, $p = .05$), and engaging in these prolonged drinking bouts within the last year (OR = 2.56, $p = .03$). Face-to-face estimates of the latter behavior during the previous 30 days are twice as high as telephone estimates, but absolute numbers are too small to draw strong conclusions. Admission of alcohol use and binge drinking during the last year are also higher among face-to-face respondents, though not significant at the .05 level (OR = 1.35, $p = .10$ and OR = 1.36, $p = .11$, respectively). Finally, those in the

Table 2
Prevalence estimates of tobacco, alcohol, and marijuana use by interview method and odds of admitting use in face-to-face interview ($n = 163$) relative to telephone interview ($n = 853$)

Substance	Estimated prevalence		OR (95% CI)	p-Value	Adj. OR (95% CI)	p-Value
	Phone	Face				
Cigarettes						
Ever	82.9	82.8	1.00 (.64–1.55)	.985	1.04 (.65–1.66)	.863
Last 30 days	57.8	60.1	1.10 (.78–1.55)	.581	1.19 (.83–1.70)	.345
Last year	65.4	65.6	1.01 (.71–1.44)	.955	1.07 (.074–1.54)	.730
Alcohol						
Ever	92.1	91.4	.91 (.50–1.66)	.751	.83 (.45–1.54)	.562
Last 30 days	30.9	39.9	1.48 (1.05–2.09)	.027	1.48 (1.03–2.11)	.033
Last year	56.9	64.4	1.37 (.97–1.34)	.076	1.35 (.94–1.94)	.104
Binge drink						
Ever	58.9	53.4	.80 (.57–1.12)	.195	.79 (.55–1.14)	.207
Last 30 days	9.7	12.3	1.30 (.77–2.18)	.326	1.36 (.80–2.32)	.255
Last year	24.6	31.3	1.39 (.97–2.01)	.075	1.36 (.93–1.99)	.112
Binge						
Ever	12.5	17.8	1.51 (.96–2.37)	.073	1.60 (1.00–2.58)	.051
Last 30 days	.6	1.2	—	—	—	—
Last year	2.2	6.1	2.87 (1.31–6.29)	.008	2.56 (1.12–5.88)	.027
Problem	17.1	16.6	.96 (.61–1.51)	.864	.99 (.62–1.59)	.977
Marijuana						
Ever	63.0	63.2	1.01 (.71–1.43)	.954	1.02 (.71–1.46)	.924
Last 30 days	5.0	10.4	2.19 (1.29–3.95)	.009	1.91 (1.03–3.54)	.040
Last year	12.8	22.7	2.00 (1.32–3.04)	.001	1.81 (1.16–2.83)	.009
Problem	11.3	9.8	.86 (.49–1.50)	.591	.86 (.48–1.54)	.611

"Binge drinking" is operationalized as five or more drinks (four for women) in one sitting. "Binge" refers to drinking for at least 2 days without sobering up. "Problem" is based upon respondent's self-report of belief they ever thought they might have had a problem with the substance. Those categories with dashes do not have enough affirmative responses in one or both surveys to reliably estimate odds ratios.

face-to-face sample are significantly more likely to admit using marijuana during the last month (OR = 1.91, $p = .04$) and last year (OR = 1.81, $p = .01$). Thus there is no difference between our telephone and face-to-face surveys of welfare recipients in terms of admitting *ever* using alcohol or marijuana or *ever* thinking one might have had a problem with them. Relative to telephone administration, however, the face-to-face survey provides higher prevalence estimates of *recent* use of marijuana and alcohol, including estimates of abuse such as multiple-day drinking episodes.

Table 3 compares estimates for hard drugs. The findings show that face-to-face interviews reveal higher estimates of recent substance use, though in most cases these differences are non-significant. The one exception is that face-to-face respondents are 4.5 times more likely to admit using a sedative[2] within the last month (OR = 4.59, $p < .01$). There is one major difference in the results for hard drugs compared to those for cigarettes, alcohol, and marijuana. For cocaine/heroine and stimulants, estimates of lifetime prevalence (OR = .62, $p = .06$, OR = .58, $p = .06$, respectively) and ever thinking that one might have had a drug problem (OR = .42, $p = .05$, OR = .48, $p = .08$, respectively) were significantly lower (or nearly so) in the face-to-face surveys. The findings for all of these drugs combined (i.e., the "any drug" category) appear to summarize the general findings for hard drugs. The face-to-face respondents are significantly less likely to admit ever using a drug or ever thinking they might have had a problem, but significantly more likely to admit use within the last month (though the latter is largely due to sedative use).

Discussion

Research on survey methodology has serious implications for public health. Better methods can provide more reliable estimates that can increase the detection of health threats and aid in more efficiently identifying the type and amount of resources required to address the

[2]In this survey, sedatives prescribed by a physician are not included in this category, though over-the-counter sedatives are included.

Table 3
Prevalence estimates of substance use by interview method and odds of admitting use in face-to-face interview ($n = 163$) relative to telephone interview ($n = 853$)

Substance	Estimated prevalence		OR (95% CI)	*p*-Value	Adj. OR (95% CI)	*p*-Value
	Phone	Face				
Cocaine/heroin						
Ever	21.3	13.5	.58 (.40–.93)	.025	.62 (.37–1.01)	.055
Last 30 days	.1	2.5	—	—	—	—
Last year	2.6	3.1	1.20 (.45–3.22)	.715	1.31 (.48–3.58)	.598
Problem	8.9	3.7	.40 (.17–.92)	.032	.42 (.18–1.00)	.049
Sedatives						
Ever	16.5	11.7	.67 (.40–1.11)	.120	.69 (.41–1.16)	.158
Last 30 days	1.3	4.9	3.95 (1.56–9.97)	.004	4.59 (1.76–11.54)	.002
Last year	5.6	6.7	1.21 (.62–2.39)	.575	1.23 (.61–2.48)	.555
Problem	3.9	3.1	.79 (.30–2.05)	.622	.85 (.32–2.24)	.738
Stimulants						
Ever	17.0	9.8	.54 (.31–.92)	.025	.58 (.33–1.03)	.063
Last 30 days	.8	1.8	—	—	—	—
Last year	4.5	3.1	.68 (.27–1.76)	.431	.76 (.30–2.00)	.580
Problem	9.3	4.3	.44 (.20–.98)	.044	.48 (.21–1.08)	.076
Hallucinogens						
Ever	13.5	9.2	.66 (.37–1.15)	.143	.70 (.39–1.25)	.227
Inject						
Ever	8.4	5.5	.64 (.31–1.30)	.218	.69 (.33–1.44)	.321
Any drug						
Ever	33.6	23.9	.62 (.42–.91)	.016	.65 (.44–.97)	.037
Last 30 days	2.0	6.1	3.21 (1.44–7.15)	.004	3.49 (1.53–7.92)	.003
Last year	9.7	9.2	.94 (.53–1.68)	.834	.97 (.53–1.75)	.908
Problem	15.5	8.6	.51 (.29–.92)	.024	.56 (.31–1.01)	.052

The "cocaine/heroin" category includes powder cocaine, crack cocaine, and heroin. The "sedatives" category includes pain relievers, tranquilizers, and sedatives. The "stimulants" category includes methamphetamine and stimulants. "Inject" means that respondent has ever injected any drug to get high. "Any drug" includes the cocaine, hallucinogens, sedatives, stimulants, and inject categories. "Problem" is based upon respondent's self-report of belief they ever thought they might have had a problem with the substance. Those categories with dashes do not have enough affirmative responses in one or both surveys to reliably estimate odds ratios.

problem. Recent studies reveal the benefits of both new technology (e.g., A-CASI) and traditional low-tech modes (e.g., paper and pencil SAQs) in reducing response bias in surveys of sensitive behavior. Tourangeau and Smith (1996), for example, show that A-CASI provides higher estimates than CAPI and CASI of illicit drug use, and Aquilino (1994) shows that SAQs provide higher rates of alcohol and drug use than telephone and face-to-face surveys. Other studies reveal higher reporting of adolescent sexual behavior, drug use, and violence using A-CASI compared to SAQs (Turner et al., 1998), higher admission with A-CASI compared to face-to-face of stigmatized behaviors such as being HIV-positive or recent use of non-prescription methadone (Newman et al., 2002; see also Perlis et al., 2004), higher estimates of alcohol-related health, work, and "any" harm in telephone compared to face-to-face methodology (Midanik, Greenfield, & Rogers, 2001), and increased self-disclosure of intimate partner violent offending and victimization via confidential computer screening in hospital emergency departments (Rhodes et al., 2002).

Not all of the findings from these recent studies are consistent, however, and our results reveal the complexity of carrying out surveys of sensitive behavior. For example, face-to-face interviews may allow the respondent to identify, judge, and build a rapport with the interviewer, which can make some respondents feel more confident in offering personal information (Martin, 1983). Alternatively, the perceived distance and anonymity of telephone surveys may have similar effects on other respondents. Since our survey is of TANF recipients, we believed respondents may be more trustful that their revelations will not be used against them if they are interviewed in person, and our results provide partial support of this hypothesis. First, when

asked to rate the quality of the data obtained in the interview given the respondent's demeanor, our interviewers employing face-to-face methodology were significantly more likely than the telephone interviewers to rate the data quality as excellent (85% and 72%, respectively, $p<.01$). Second, although telephone respondents were as willing as face-to-face respondents to report *ever* using or having a problem with alcohol or marijuana and more willing to report *ever* using or having a problem with harder drugs such as cocaine and stimulants, self-reports of *recent* drug or alcohol abuse were higher in the face-to-face survey. This contrasts with some earlier research and may hint at the differential impact of mode of administration by the type of population under study, which is also evident in the work of Greenfield et al. (2000) and Aquilino and LoSciuto (1990), who show that mode effects may vary by income and race, respectively.

Ever using a drug or ever thinking you might ever have had a problem with it is not as helpful in discovering current abuse and need for treatment as are questions about recent use. Since our face-to-face interviews reveal higher estimates of recent use (and abuse in terms of the alcohol questions) this mode may be more appropriate, especially for this population and especially if accompanied by specific questions about drug- and alcohol-related problem behavior and abuse. The Newman et al. (2002) results are relevant to this notion, since they show that face-to-face interviews yield higher estimates of psychological distress. For example, they found that estimates of recent (i.e., at least two-thirds of the time within the last month) feelings of hopelessness and depression were significantly higher among face-to-face respondents (compared to A-CASI). Recent (ab)use may indicate another type of distress and we believe our findings help to broaden those of the Newman et al. study.

Limitations

There are a few main limitations to this study. First, we have no way to validate self-reports of alcohol and drug use. Recent research using ADAM data (National Institute of Justice, 2003) has compared self-reported drug use with urinalysis results and the findings suggest that under-reporting is less likely when respondents perceive the risk of discovery to be higher. While this sort of methodology is impossible with most self-report surveys, the NIJ results appear to confirm the standard assumption of studies such as ours that higher prevalence estimates of sensitive behavior are indicative of more accurate reporting. Second, we have essentially an all female sample, and women may be better reporters and observers of their health status than males and may be more likely to respond to requests to be surveyed. While this may limit the generalizability of our results, this demographic makeup is consistent with TANF recipients as a whole and thus our results have practical relevance. Third, the number of respondents in our face-to-face survey was too small to obtain reliable estimates of some rare events such as injection and recent use of cocaine and stimulants. These are dangerous substances and recent use may signal need for treatment. Thus our study would have benefitted from a larger face-to-face sample given the importance of this information.

Fourth, since this is a within-group comparison of TANF recipients we are unable to discern if our distinct pattern of results is due to caution specific to our TANF population or to the differential response rates typical of these survey modes. Future work should therefore compare between-group estimates of this group (or some similar group of interest) and the overall population. Nevertheless, our main interest is a practical one and our results have important implications since we are concerned precisely with mode effects on this specific population of welfare recipients who may feel they have something very tangible to lose based upon how they respond to questions about their drug use. Thus despite the potentially confounding effects, our findings will be beneficial to local and state health departments and other agencies concerned with obtaining reliable estimates of these behaviors from similar populations. Fifth, the effect of performing the face-to-face surveys using CAPI versus PAPI (paper and pen) in this study is unclear. According to Baker (1992), for example, CAPI administration may result in reduced rapport because of the intrusiveness of the equipment, the delay between questions, and the lack of eye contact compared with PAPI administration. This weakened rapport is thought to decrease the willingness to disclose sensitive information. Still, the fact that we found significant differences between the modalities attests to the power of the face-to-face environment regardless of whether PAPI or CAPI is used.

Sixth, we recognize that we have not created a true experimental condition, where data are collected using different modalities from two randomly assigned samples. In fact, it may be that the reason our face-to-face CAPI respondents are not in our telephone CATI sample is because they are "different" from the CATI sample. Our data show, for example, that the face-to-face CAPI sample is more likely to be employed at least part time and are more likely to have ever been married than the CATI sample. It is not clear, though, how work and marital status might influence the willingness of a respondent to disclose sensitive information. Indeed, one might argue that respondents who are employed or married would be *less* willing to disclose. In any event, we have controlled the differences between the samples by calculating adjusted odds ratios, which are reported in Tables 2 and 3.

Finally, our sample represents a unique group and our findings may not generalize to the overall population.

Given the evaluative nature of our study the results are important regardless of their representativeness of the overall population. This is especially true since our population represents a group that has not been systematically examined by this type of research, even though they have high rates of drug use and limited personal resources and thus often are dependent on state agencies for treatment. Nevertheless, we must point out that neither the estimates nor the pattern of mode effects may be generalizable beyond our population.

Conclusion

Some researchers have called for increased methodological research on mode effects for specific groups such as minorities (Jackson & Ivanoff, 1999), and others have pointed out differential prevalence estimates by income level (Greenfield et al., 2000) and race (Aquilino & LoSciuto, 1990). Further, tens of millions of dollars are spent annually by governmental and public health agencies that target specific groups, such as TANF recipients, when planning and providing treatment and other responses for alcohol and substance abuse and related problems. We know that the level of respondents' trust can vary substantially by age, minority status, and income level, and also by the type of survey mode (Wright, Aquilino, & Supple, 1998). This variation may significantly influence survey non-response and patterns of reporting sensitive behavior, so it is vital that we undertake research on survey methodology for these specific groups in order to understand how to obtain trustworthy estimates upon which to base an efficient and effective public health response. This is especially important when we remember that a random sample does not assure valid results in the face of the substantial and multiple threats of non-sampling error (Assael & Keon, 1982).

We found that if a researcher is simply interested in *lifetime* prevalence or in whether the respondent believes s/he might *ever* have had a drug or alcohol problem, telephone administration appears to be superior, since in spite of higher non-response rates prevalence estimates are equal to or higher than those of face-to-face surveys after adjusting for any differences between the samples. Thus equally dependable responses about some types of sensitive behavior can be obtained from the respondent at a lower cost (McAuliffe, Geller, LaBrie, Paletz, & Fournier, 1998; Tyebjee, 1979). In this study, for example, the face-to-face CAPI sample had the increased cost of interviewer travel and salary for supervisors and interviewers. Salary for supervisors was compounded by the lengthened period of data collection (3 months for face-to-face CAPI compared to 1 month for the telephone CATI survey). Including all costs, the cost of each completed face-to-face CAPI survey was $238. The cost of each completed telephone CATI survey was $96. On the other hand, if the agency or researcher requires information about *recent* use and potential abuse (and thus about increasingly sensitive information that the respondent may be less willing to reveal), face-to-face interviews may provide more accurate information.

It is necessary, however, to examine more specific questions related to substance abuse and self-perceived treatment needs before drawing strong conclusions. Further research is required with specific populations targeted by agencies, since our results (from a specific and important group) differ somewhat from those of other studies (that generally examine samples representative of the overall population). This suggests that we cannot assume that knowledge of mode effects and response bias about the larger population applies to specific sub-groups. Future research should focus on (1) between-group (e.g., overall population versus welfare population) comparisons of estimates for individual modes to determine if differences exist, (2) specific populations often studied for public health and scientific purposes, (3) the efficacy of other modes (e.g., A-CASI) with these groups, and (4) more specific questions about abuse and need for treatment. Since it appears that the selection of survey mode might depend not only on the type of information required but also on the characteristics of the population under study, such research could have a profound effect on the validity of the information we are able to gather and thus will have an equally important influence on efficiently planning for and responding to individual and public health needs.

Acknowledgements

The data collection for this research was supported by Contract No. 270-98-7066 awarded by the Center for Substance Abuse Treatment to the Oklahoma Department of Mental Health and Substance Abuse Services. Points of view are those of the authors and do not necessarily represent the official positions of CSAT or ODMHSAS. The authors thank Lior Gideon and Colin Loftin for their helpful comments on and critiques of an earlier draft. The first author thanks the Davis Center at Harvard University, where he was a Research Fellow when this article was written. An earlier version of this paper was presented at the 2002 annual meeting of the American Society of Criminology, Chicago.

References

Aquilino, W. S. (1994). Interview mode effects in surveys of drug and alcohol use: A field experiment. *The Public Opinion Quarterly*, *58*, 210–240.

Aquilino, W. S., & LoSciuto, L. A. (1990). Effects of interview mode on self-reported drug use. *The Public Opinion Quarterly*, *54*, 362–395.

Aquilino, W. S., & Wright, D. L. (1996). Substance use estimates from RDD and area probability samples: Impact of differential screening methods and unit nonresponse. *The Public Opinion Quarterly*, *60*, 563–573.

Assael, H., & Keon, J. (1982). Nonsampling vs. sampling errors in survey research. *Journal of Marketing*, *46*, 114–123.

Baker, M. P. (1992). New technology in survey research: Computer assisted personal interviewing (CAPI). *Social Science Computer Review*, *10*, 145–157.

CASRO. (2002). Available online at http://home.clara.net/sisa/resprhlp.htm.

Couper, M. P., & Rowe, B. (1996). Evaluation of a computer-assisted self-interview component in a computer-assisted personal interview survey. *The Public Opinion Quarterly*, *60*, 89–105.

Crook, G. M., & Oei, T. P. S. (1998). A review of systematic and quantifiable methods of estimating the needs of a community for alcohol treatment services. *Journal of Substance Abuse and Treatment*, *15*, 357–365.

Czaja, R. (1987). Asking sensitive behavioral questions in telephone interviews. *International Quarterly of Community Health Education*, *8*, 23–32.

DeMaio, T. J. (1984). Social desirability and survey measurement: A review. In C. F. Turner, & E. Martin (Eds.), *Surveying subjective phenomena* (pp. 257–281). New York: Russell Sage.

Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, NY: Doubleday/Anchor Books.

Greenfield, T. K., Midanik, L. T., & Rogers, J. D. (2000). Effects of telephone versus face-to-face interview modes on reports of alcohol consumption. *Addiction*, *95*, 277–284.

Henson, R., Cannell, C. F., & Roth, A. (1978). Effect of interview mode on reporting of moods, symptoms, and need for social approval. *Journal of Social Psychology*, *105*, 123–129.

Herzog, A. R., & Rodgers, W. L. (1988). Interviewing older adults. Mode comparison using data from a face-to-face survey and a telephone resurvey. *The Public Opinion Quarterly*, *52*, 84–99.

Jackson, A. P., & Ivanoff, A. (1999). Reduction of low response rates in interview surveys of poor African–American families. *Journal of Social Science Research*, *25*, 41–60.

Jones, E. F., & Forrest, J. D. (1992). Underreporting of abortion in surveys of US women: 1976–1988. *Demography*, *29*, 113–126.

Johnson, T. P., Hougland, J. G., & Clayton, R. R. (1989). Obtaining reports of sensitive behavior: A comparison of substance use reports from telephone and face-to-face interviews. *Social Science Quarterly*, *70*, 174–183.

Lebkowski, J. M. (1988). Telephone sampling methods in the United States. In R. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholis, & J. Waksberg (Eds.), *Telephone survey methodology*. New York: Wiley.

Mangione, T. W., Hingson, R., & Barrett, J. (1982). Collecting sensitive data: A comparison of three survey strategies. *Sociological Methods and Research*, *10*, 337–346.

Martin, E. (1983). Surveys as social indicators: Problems in monitoring trends. In P. H. Rossi, J. D. Wright, & A. B. Anderson (Eds.), *Handbook of survey research* (pp. 677–743). Orlando, FL: Academic Press.

McAuliffe, W. E., Geller, S., LaBrie, R., Paletz, S., & Fournier, E. (1998). Are telephone surveys suitable for studying substance abuse? Cost, administration, coverage and response rate issues. *Journal of Drug Issues*, *28*, 455–481.

McQueen, D. V. (1989). Comparison of results of personal interview and telephone surveys of behavior related to risk of AIDS: Advantages of telephone techniques. In F. Fowler (Ed.), *Health survey research methods: Conference proceedings* (pp. 247–252). Washington, DC: Government Printing Office.

Midanik, L. T., Greenfield, T. K., & Rogers, J. D. (2001). Reports of alcohol-related harm: Telephone versus face-to-face interviews. *Journal of Studies on Alcohol*, *62*, 74–78.

National Institute of Justice. (2003). *2000 arrestee drug abuse monitoring: Annual report*. Washington, DC: United States Department of Justice.

Nebot, M., Celentano, D. D., Burwell, L., Davis, A., Davis, M., Polacsek, M., & Santelli, J. (1994). AIDS and behavioural risk factors in women in inner city Baltimore: A comparison of telephone and face-to-face surveys. *Journal of Epidemiology and Community Health*, *48*, 412–418.

Newman, J. C., Des Jarlais, D. C., Turner, C. F., Gribble, J., Cooley, P., & Paone, D. (2002). The differential effects of face-to-face and computer interview modes. *American Journal of Public Health*, *92*, 294–297.

Perlis, T. E., Des Jarlais, D. C., Friedman, S. R., Arasteh, K., & Turner, C. F. (2004). Audio-computeriszed self-interviewing for research data collection at drug abuse treatment programs. *Addiction*, *99*, 885–896.

Rhodes, K. V., Lauderdale, D. S., He, T., Howes, D. S., & Levinson, W. (2002). "Between me and the computer": Increased detection of intimate partner violence using a computer questionnaire. *Annals of Emergency Medicine*, *40*, 476–484.

Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *The Public Opinion Quarterly*, *60*, 275–304.

Turner, C. F., Danella, R. D., & Rogers, M. (1995). Sexual behavior in the United States, 1930–1990: Trends and methodological problems. *Sexually Transmitted Diseases*, *22*, 173–190.

Turner, C. F., Ku, L., Rogers, S. M., Lindberg, L. D., Pleck, J. H., & Sonenstein, F. L. (1998). Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science*, *280*, 867–873.

Tyebjee, T. T. (1979). Telephone survey methods: The state of the art. *Journal of Marketing*, *43*, 68–78.

US Department of Health and Human Services. (2003). Results from the 2001 National Household Survey on Drug Abuse. Available online at http://www.samhsa.gov/oas/drugs.htm.

Wright, D. L., Aquilino, W. S., & Supple, A. J. (1998). A comparison of computer-assisted and paper-and-pencil self-administered questionnaires in a survey on smoking, alcohol, and drug use. *The Public Opinion Quarterly*, *62*, 331–353.