

DO INTERVIEWER-RESPONDENT RACE EFFECTS IMPACT THE MEASUREMENT OF ILLICIT SUBSTANCE USE AND RELATED ATTITUDES?

David Livert, Charles Kadushin, CUNY Graduate Center, New York City
Mark Schulman, Andy Weiss, Schulman, Ronca, & Bucuvalas, Inc., New York City

Key Words: Interviewer Effects, Substance Use, Attitudes

Race of interviewer effects have been among the most studied sources of non-random measurement error attributable to the survey interviewer, dating back to Hyman's (1954) study of attitudes among black respondents in Memphis. He found that blacks interviewed by whites were more likely to express patriotic attitudes than those interviewed by blacks. A growing body of research demonstrates similar effects (Groves, 1989). Race of interviewer effects have been found with questions with racial content, particularly when the question refers to race of interviewer, are subject to social desirability pressures, or concern political or economic institutions (Campbell, 1981; Schaeffer, 1980). Questions about substance use may be subject to such effects to the degree to which they are impacted by social desirability motivations.

Even without the face-to-face contact of a personal interviewer, similar race effects have been found in telephone interviews (e.g., Cotter, Cohen, & Coulter, 1982; Davis, 1997). Indeed, Davis (1997) found that although the perceived race of telephone interviewers generally matched the actual race of telephone interviewers, both perceived interviewer race and actual interviewer race had equivalent impact on responses to race-oriented political questions.

Telephone interviews may provide conservative tests of race of interviewer effects: non-verbal information regarding the interviewer is not available to the respondent (Davis, 1997). Even when interviewer characteristics are perceived by the respondent, motivation to alter behavior based on this information may be lessened by the greater social distance over the phone lines. Nonetheless, it is particularly important to explore the potential for such effects in this modality, due to the much larger number of interviews completed by any one interviewer compared with personal interviews (Singer, Frankel, & Glassman, 1983).

Theoretical explanations of interviewer race effects include respondent "deference" or cautiousness, interpersonal politeness, and social distance. Older racial deference (Hyman, 1954) or more recent cautiousness (Davis, 1997) explanations suggest that minority respondents may under-report socially undesirable behavior to majority interviewers out of concerns about potential power differentials.

Asymmetric patterns of interviewer effects would be consistent with this explanation: one would expect results for blacks interviewed by whites, but not necessarily whites interviewed by blacks. A comparison of interviewer race effects across both white and black respondents led Hatchett and Shuman (1975, Anderson, Silver, & Abramson, 1988b) to propose an interpersonal politeness explanation: in cross-race interviewing situations, an interpersonal norm of not offending the interviewer may override that of task norms (i.e. providing accurate responses). For example, whites may provide more "pro-black" responses when interviewed by blacks. For sensitive questions, greater social distance in cross-race interviews may inhibit the development of rapport.

Interviewer-respondent race effects may also arise from the interviewer, rather than the respondent's reaction to the interviewer. Interviewers may be less comfortable asking sensitive questions in cross-race interviewing situations (Schaeffer, 1980). Additionally, interviewer expectations regarding the ease of asking specific items has been linked to the quality of item response (Singer, Frankel, & Glassman, 1983). Such expectations may be expressed through tone or lack of probing by the interviewer.

Mode Effects in Measurement of Substance Use

Most studies utilize individual self-reports to estimate the incidence of illicit substance use. Due to social desirability bias, such self-reports systematically underestimate substance use. However, such under-reporting tends to vary by survey mode. Generally, the highest rates of substance use reporting occurs in self-administered questionnaire modalities (SAQ) or personal interviews which integrates SAQ administered in a personal interview situation. Standard personal interviewing yields somewhat lower rates and telephone interviews yield the lowest (Turner, Lessler, & Gfroerer, 1992). Face-to-face interviews, in turn, may lead to greater under-reporting than self-administered modalities. Gfroerer & Hughes (1991) reported that the NHSDA, which used self-administered sheets for alcohol and drugs, yielded higher estimated rates of substance use than a national telephone survey utilizing comparable questions. Turner, Lessler, and Devore (1992) found higher rates for marijuana and cocaine use in self-administered versions of the NHSDA compared with face-to-face interviewer administration. Fendrich and Vaughn (1994) found less reporting of marijuana and

cocaine use in telephone follow-ups of the NLSY than in face-to-face interviews. Likewise, Aquilino (1994) found lower estimates of crack, cocaine, use, and marijuana in telephone surveys when compared with personal interviews using a SAQ. Higher rates of illicit drug use and sexual behavior have been reported using audio-CASI as compared to face-to-face interviewing (Turner et al., 1998). Respondents may be more likely to under-report in telephone interviews because the lack of non-verbal interaction attenuates respondents' comfort that his or her responses will be held in confidence (Aquilino, 1994; Groves, 1990), and because telephone interviews fail to provide respondents with sufficient cues to interviewer identity that might encourage self-report of socially undesirable behaviors (Johnson, Houghland & Moore (1991).

Substance Use Reporting and Respondent Race

Several studies suggest that substance use under-reporting and inconsistency may vary by respondent race. Mensch and Kandel's (1988) study of the NLSY revealed greater inconsistencies in the reporting of marijuana use among blacks and Hispanics than among white respondents. Fendrich & Vaughn (1994) also found lower rates and greater inconsistencies in reporting of marijuana and cocaine use among black and Hispanic respondents than among white and other non-white respondents.

Survey mode effects on under-reporting also vary by respondent race. For alcohol, marijuana, cocaine, and crack (Aquilino, 1994), differences between survey methodologies (telephone, personal with SAQ, personal without SAQ) were larger for blacks than whites; there were ambiguous findings for Hispanics (due to small sample sizes). Mode effects on under-reporting were also larger for respondents who indicated greater mistrust of the interviewing process, although controlling for trust did not eliminate mode/race differences. Fendrich and Vaughn (1994) found that blacks' under-reporting of marijuana and cocaine use was consistent across interview modes; Hispanics were more likely to under-report use in a SAQ than a face-to-face mode.

Race effects in under-reporting and differential under-reporting across modes may be due to variations in social distance, status inequalities (Fendrich & Vaughn, 1994), or a general motivation on the part of minority groups to provide more socially desirable responses in such surveys (Mensch & Kandel, 1988).

Fendrich and Vaughn (1994) have suggested that differential under-reporting among minority respondents needs to be examined in greater detail; one such line of inquiry should investigate effects between minority respondents and interviewers. To our knowledge, few, if any, studies have examined such effects on substance use reports over the telephone.

Related research on mode effects is inconclusive. Analyzing the face-to-face NLSY, Mensch & Kandel (1988) failed to find effects of interviewer race or gender on inconsistencies in reporting marijuana use. Fendrich and Vaughn (1994) failed to find interviewer effects in the NLSY administered in three survey modes: self-report, personal interview, and telephone. Aquilino (1994) found that race of interviewer impacted the differences between personal interview and personal interview/SAQ for Hispanic respondents on alcohol items but not for black respondents; however, direct effects of interview-respondent configuration were not reported.

Rationale for Study

Given the growing concern regarding mode effects on substance use reporting and the potential for such effects to vary by race of respondent, we examined whether the configuration of respondent and interviewer race impacted the reporting of use in a large telephone survey. We assume the misuse of alcohol and the use of illicit substances are sensitive items vulnerable to social desirability. Attitudes and beliefs about substance use may not represent sensitive questions (at least to the degree that self reports of such use would), but are subjective and may be equally susceptible (Tucker, 1983). Both classes of questions provide key data regarding the incidence of substance use and the effectiveness of programs designed to prevent abuse. For comparative purposes, we examined a set of questions not related to AOD use.

If detected, interviewer-respondent race effects may fit either of two patterns. Such effects could be due to perceived social distance with cross-race configurations producing greater distance and thus lower certainty of confidentiality. Greater rates -- assumed to be more accurate (Miller, 1998) -- should be evidenced in same-race combinations of interviewer and respondent, regardless of the race/ethnicity. This pattern of results would also be consistent with an interpersonal politeness explanation. Alternatively, race effects may be due to cautiousness arising from minority-majority status of interviewers: Minority respondents may feel more threatened by white interviewers asking questions about illicit drug use than non-whites. Interviewer-respondent effects limited to non-white respondents would support this explanation.

This survey data set offers sufficient power to explore these effects with all configurations of white and black interviewers and respondents; there are also sufficient numbers of Hispanic respondents in the study to investigate whether such effects exist among Hispanic respondents interviewed in English by white and black interviewers.

Method

This study is based on data collected through

telephone interviews with residents aged 16-44 in 41 mid-sized urban communities across the United States. The data were collected as part of an ongoing national evaluation of the Robert Wood Johnson Foundation's Fighting Back initiative to reduce substance use and related harm. Concentrated on urban communities, the survey sample has a higher incidence of minority respondents than general population surveys of the U.S. Interviews were conducted by SRBI's centralized telephone facility in New York City, providing a diverse telephone interviewer staff. Taken together, this combination of diverse respondents and interviewers provides sufficient power to examine interviewer-respondent effects on substance use reporting.

Sample

Data are drawn from Wave II of the Fighting Back evaluation survey, which consisted of 17,900 interviews conducted in 41 communities during the spring of 1997. The profile of respondents reflects its urban concentration. Less than half (41%) of respondents are white; 34% were black. Although Hispanic ethnicity was measured by a separate question, for simplification any respondent answering Hispanic was reclassified from white, black or other racial classification to Hispanic, which then constituted 19% of the sample. The sampling frame focused on ages 16 to 44, among whom substance use activity is most prevalent; roughly a sixth (17%) of the sample was between 16 and 20. Forty-one percent of respondents were in the South, 25% in the Northeast and the remainder in the Midwest and West census regions.

Interviewer characteristics consisted of gender, race, and number of completed interviews. Over half (58%) of interviewers were female; half were black (50%) and another 42% were white. The median number of interviews conducted by a single interviewer was 29 and ranged from 1 to 273.

To provide sufficient power for interviewer-respondent effects, analysis of interviewer effects were restricted to white and black interviewers and white, black and Hispanic respondents, yielding a sample of 12,872 respondents crossed with 343 interviewers. By design, the interviews were interpenetrated: theoretically, respondents were randomly assigned to each interviewer. Examination of assignment of respondent characteristics to black and white interviewers suggests that complete random assignment was not attained.

Interviewer gender varied by interviewer race: there were more black female interviewers than white female interviewers. Workload did not vary by race of interviewer. Respondent gender, age, and region of interview as well as interviewer gender and interview load were included as covariates to control for potential differences in interviewer assignment. However, an

assumption of random assignment may not be credible with regard to interviewer and respondent race; we cannot rule out the possibility that this source of non-random error may have influenced the findings presented herein.

Variables Examined

Three classes of dichotomous variables in the Wave II questionnaire were examined. Substance use items included current tobacco smoking, 30 day alcohol use, alcohol bingeing (5 or more drinks at one time) during the past month, 12 month, 30 day, and weekly marijuana use, 12 month cocaine use, and 12 month barbiturate use. Substance use attitudes included five items measuring the risk of harm from using alcohol, marijuana, cocaine, crack, and heroin as well as three items indicating personal disapproval of bingeing, marijuana and crack. The risk and disapproval items were dichotomized as follows: "great risk" vs. all other categories and "strongly disapprove" vs. all other categories. Five other items were also included for comparison purposes: whether the respondent lived in the same neighborhood for over a year, involvement in volunteer activities, current health status, rating of crime as a serious neighborhood problem, and whether someone had broken into the respondent's household in the past 12 months.

Results

The hierarchical nature of the survey data -- respondents nested within interviewers -- requires statistical techniques appropriate for the assessment of clustered data, that may not be sufficiently controlled for in individual-level statistical analysis which assume independence of observations in the data. The result can be a serious underestimation of standard errors for interviewer-level effects, which may misrepresent interviewer effects (Bryk & Raudenbush, 1992; Goldstein, 1995). Thus, effects were tested using the MLN (Woodhouse, 1995) multilevel modeling program.

Effects were tested using two-level logistic regression. The statistical model consisted of dummy codes composed of respondent characteristics, interviewer characteristics, and their interactions. Respondent characteristics consisted of gender (male = contrast category), race (white = contrast), respondent age (21-44 = contrast), and census region (northeast = contrast). Interviewer characteristics were gender (male = contrast) and race (white = contrast). A relationship between number of interviews completed and interviewer effects might arise from variations in interviewer load: those who interviewed only a few respondents might not have learned the questionnaire. On the other hand, effects might be due to interviewer carelessness or boredom arising from an above average number of completed surveys. To test for these effects, the number

of completed interviews was coded into a trichotomous variable for each interviewer (1 - 9; 10 - 74; 75 or more). To test the effects due to configuration of interview and respondent race, two dummy codes representing their interaction were entered last and the associated improvement in model fit was subjected to a Chi Square test.

Impact of Interviewer Characteristics

The primary focus of this study was whether respondents varied their reporting of substance use and related attitudes in reaction to the race of the telephone interviewer; it is tested statistically by examining the terms representing the interaction of respondent and interviewer race. However, before turning to this interaction, main effects of interviewer characteristics were examined. Controlling for respondent characteristics (sex, age and region), interviewer gender had no significant effects on the 21 outcomes. Although black interviewers elicited somewhat higher rates of 12 month marijuana use than did white interviewers, no other main effects for interviewer race on substance use items were significant. Perceived harm from crack and heroin use reported to black interviewers was higher than for white interviewers; on the other hand, disapproval of bingeing and marijuana use was lower. Interviewers completing more than 75 interviews were likely to have somewhat lower rates of perceived harm from crack and heroin and lower rates of disapproval of crack use. One interviewer effect was detected for non-substance use items: respondents were more likely to report living in the same neighborhood for longer than a year to white interviewers.

Interaction between Race of Respondent and Race of Interviewer

Significant respondent-interviewer race effects emerge for alcohol and marijuana use when considering specific interviewer-respondent combinations. To interpret the nature of these interactions, significant comparisons and accompanying estimates of rates are presented in Table 1. Black respondents are more likely to report 30 day alcohol use and bingeing to black interviewers than white interviewers. Rates for alcohol use among white and Hispanic respondents did not vary by race of interviewer. Both black and Hispanic respondents were more likely to report marijuana use — either 12 month, monthly, or weekly — to black interviewers than white interviewers. No interviewer-respondent effects were found for 12 month cocaine and barbiturate use.

Interviewer effects for black and Hispanic respondents were also found for perceptions of marijuana use. Both groups were more likely to consider marijuana use harmful and to disapprove of such use when interviewed by whites. In addition, Hispanics were more

likely to rate bingeing as harmful and blacks were more likely to disapprove of bingeing with white interviewers. No significant effects were found for perceived harm from and disapproval of cocaine, crack and heroin. No effects were found for white respondents.

Among non-substance use outcomes, white, black, and Hispanic respondents were more likely to report volunteering in a community organization to white interviewers than black. Blacks reported poorer health to black interviewers than to whites.

Further analysis tested whether the pattern of interviewer effects was conditioned by respondent sex, interviewer age, and region of interview. Examination of these three-way interactions generally did not alter our findings. Reports of bingeing was moderated by region: Blacks in the South reported higher rates to black interviewers than to whites, this difference was not significant for the other three regions. Interviewer race effects for perceived harm from bingeing was limited to Hispanic males. Greater interviewer-respondent effects were found for 12 month marijuana use for blacks and Hispanics interviewed by females compared with those interviewed by males. Effects for weekly marijuana use were also weaker for Hispanics residing in the South than those residing elsewhere.

Impact of Interviewer Effects on Estimated Rates of Outcomes

We were interested to know whether our estimates of substance use and related attitudes for whites, blacks, and Hispanics in the survey would have differed if a congruent combination of interviewer and respondent race had been utilized: whites interviewed by whites, blacks interviewed by blacks, and -- given the pattern yielding the higher rates -- Hispanic respondents interviewed by black interviewers. Estimates from these combinations were compared with the 95% confidence interval estimates for the 21 outcomes that were generated without consideration of interviewer characteristics.

In all but three cases, the estimates derived from a "congruent" combinations of respondent and interviewer fell within the 95% confidence interval estimates of outcomes for whites, blacks, and Hispanics. However, three out of 21 items is greater than would be expected by chance (14%) suggesting that estimates of the perceived harm from substance use may be more vulnerable to interviewer-respondent race effects. The three exceptions include binge disapproval for blacks and marijuana disapproval for black and Hispanic respondents. In these cases, a congruent configuration would have yielded lower estimates of substance use disapproval.

Conclusions

We have identified significant interviewer-

respondent race effects in a telephone survey measurement of drug use and drug use attitudes. There is a general pattern to these effects: black and Hispanics who are interviewed by whites are less likely to report substance use and more likely report disapproval and perceived harm from such use.

The results are not equally distributed by substance or by race of respondent. Effects were found for alcohol items for black respondents and for marijuana items for black and Hispanic respondents. No effects were detected for cocaine and barbiturates, although the low incidence and restricted variance associated with these substances may have prevented the detection of an effect.

We are unsure whether the restriction of interviewer-respondent effects to black and Hispanic respondents is due to differential motivations for social desirability or wariness with white interviewers. Perhaps a white interviewer asking questions about drug use and drug use perceptions may raise concerns regarding confidentiality. Alternatively, for non-whites, white interviewers may make salient a respondent's own race or ethnicity and motivate respondents to give more socially desirable responses out of group presentation concerns.

Slight differences in findings between Hispanics and blacks may be due to the lack of Hispanic-Hispanic configurations of interviewer and respondent or the relatively smaller pool of Hispanic respondents. However, differences in patterns may also be due to cultural variations in the perceived sensitivity of a particular question (Aquilino, 1994); whether substance use questions systematically vary in sensitivity by respondent race and ethnicity remains an important topic for inquiry.

Furthermore, we have assumed that the more socially desirable responses given by respondents in cross-race interviewing situations are the ones that are less "accurate." Given demonstrated under-reporting of substance use, we too adopt the position that "more is better" or closer to accuracy. However, we agree with Miller (1998) that what is socially desirable is culturally and contextually dependent. An alternative perspective on our findings is that black respondents, for example, over report marijuana use to black interviewers. Further research is clearly required to better articulate how item sensitivity and social desirability play out between groups.

Given the influence of interviewer attitudes and expectations on item non-response and response quality (Singer et al., 1983), it may be useful to explore whether interviewers with different characteristics have different expectations regarding questions and respondent performance which may impact the interview. Could it

be that these interviewer race effects are due to differences in these expectations rather than differences in social distance that arise from various combinations of interviewer and respondent?

It is possible that the effects observed in this study account for some of the consistently low reports of substance use among blacks. Such effects did not alter estimates for whites, blacks, and Hispanics because of a fairly even split between white and black interviewers conducting the survey. In studies where there is a real possibility that blacks will be disproportionately interviewed by whites, there may be greater opportunities for under-reporting arising from interviewer and respondent race combinations to occur.

References

- Anderson, B.A., Silver, B.D., & Abramson, P.R. (1988). The effects of the race of the interviewer on race-related attitudes of black respondents in SRC/CPS national election studies. Public Opinion Quarterly, *52*, 289-324.
- Aquilino, W.S. (1994). Interview mode effects in surveys of drug and alcohol use: A field experiment. Public Opinion Quarterly, *58*, 210-240.
- Bryk, A.S., & Raudenbush, S.W. (1992). Hierarchical Linear Models: Applications and Data Analysis Methods. Newbury Park, CA: Sage.
- Campbell, B.A. (1981). Race-of-interviewer effect among southern adolescents. Public Opinion Quarterly, *45*, 231-235.
- Cotter, P.R., Cohen, J., & Coulter, P.B. (1982). Race-in-interviewer effects in telephone interviews. Public Opinion Quarterly, *46*, 278-284.
- Davis, D.W. (1997). Nonrandom measurement error and race of interviewer effects among african americans. Public Opinion Quarterly, *61*, 183-207.
- Fendrich, M., & Vaughn, C.M. (1994). Diminished lifetime substance use over time: An inquiry into differential underreporting. Public Opinion Quarterly, *58*, 96-123.
- Goldstein, H. (1995). Multilevel Statistical Models (2nd ed.). New York: Halstead Press.
- Grfoerer, J.C., & Hughes, A.L. (1991). The feasibility of collecting drug abuse data by telephone. Public Health Report, *106*, 384-393.
- Groves, R.M. (1989). Survey errors and survey costs. New York: Wiley.
- Groves, R.M. (1990). Theories and methods of telephone surveys. Annual Review of Sociology, *16*, 221-240.
- Hatchett, S., & Schuman, H. (1975). White respondents and race-of-interviewer effects. Public Opinion Quarterly, *39*, 523-528.
- Hyman, H.H., Cobb, W.J., Feldman, J.J., Hart, C.W., & Spencer, C.H. (1954). Interviewing in social

research. Chicago: University of Chicago Press.

Johnson, T.P., Hougland, J.G., & Moore, R.W. (1991). Sex differences in reporting in sensitive behavior. *Sex Roles*, 24, 669-680.

Mensch, B.S., & Kandel, D.B. (1988). Underreporting of substance use in a national longitudinal youth cohort: Individual and interviewer effects. *Public Opinion Quarterly*, 52, 100-124.

Miller, P. (1998). Review: Is "up" right? The National Household Survey on drug abuse. *Public Opinion Quarterly*, 61, 627-641.

Schaeffer, N.C. (1980). Evaluating race-of-interviewer effects in a national survey. *Sociological Methods and Research*, 8, 400-419.

Singer, E., Frankel, M.R., & Glassman, M.B. (1983). The effect of interviewer characteristics and expectations on response. *Public Opinion Quarterly*, 47, 68-83.

Tucker, C. (1983). Interviewer effects in telephone surveys. *Public Opinion Quarterly*, 47, pp. 84-95.

Turner, C.F., Lessler, J.T., & Devore, J. (1992). Effects of mode of administration and wording on reporting of drug use. In C. F. Turner, J. T. Lessler, & J. C. Gfroerer, J.C. *Survey measurement of drug use: Methodological studies* (pp. 177-220). (DHHS Publication No. ADM 92-1929). Washington, DC: Government Printing Office.

Turner, C.F., Ku, L., Rogers, S.M., Lindberg, L.D., Pleck, J.H., & Sonenstein, F.L. (in press). Adolescent sexual behavior, drug use, and violence: New survey technology detects elevated prevalence. *Science*.

Woodhouse, G. (Ed.). (1995). *A Guide to MLN for New Users*. London: Insititute of Education, University of London.

Table 1: Interviewer-Respondent Race Effects: Estimated Rates of Outcome Variables

	White Respondents		Black Respondents		Hispanic Respondents	
	White Int	Black Int	White Int	Black Int	White Int	Black Int
<u>Substance Use</u>						
CURRSMOK	26.0%	24.5%	17.9%	18.0%	16.4%	17.6%
ALCH30	76.9%	75.4%	55.9%	< 59.0%	63.2%	65.1%
BINGE	36.7%	35.4%	16.4%	< 18.6%	30.8%	31.1%
MJ12	14.6%	14.1%	8.8%	< 11.1%	8.5%	< 10.8%
MJMONTH	8.5%	7.4%	4.9%	< 6.4%	3.7%	< 5.9%
MJWEEK	5.6%	4.9%	2.8%	< 4.8%	3.1%	< 4.4%
COCAINE	2.1%	2.3%	1.6%	1.3%	1.2%	2.5%
BARB	2.6%	3.1%	2.1%	1.5%	2.9%	2.6%
<u>Substance Use Attitudes</u>						
BINGHARM	29.3%	30.4%	48.7%	47.6%	43.5%	> 40.1%
BINGDIS	19.6%	19.9%	41.6%	> 35.1%	28.3%	27.2%
MJHARM	20.1%	21.7%	35.5%	> 32.3%	35.6%	> 32.9%
MJDIS	22.5%	22.6%	40.3%	> 33.7%	43.3%	> 36.1%
COKEHARM	58.8%	60.2%	71.3%	70.8%	65.3%	63.7%
CRAKHARM	81.8%	83.1%	87.1%	87.7%	86.5%	85.4%
CRAKDIS	75.7%	72.0%	79.4%	77.8%	77.8%	72.6%
HEROHARM	85.3%	86.3%	88.1%	88.6%	88.6%	87.7%
<u>Non-Substance Use Outcomes</u>						
SAMENEIG	80.3%	78.8%	81.7%	79.4%	82.1%	80.2%
VOLORG	37.1%	< 39.9%	34.6%	> 32.9%	33.4%	> 29.8%
SERCRIME	9.5%	9.6%	22.8%	21.0%	19.2%	18.0%
BREAK12	8.7%	9.9%	10.1%	11.1%	13.8%	14.4%
HEALTH	37.3%	37.5%	38.3%	> 33.5%	34.5%	35.0%

Note: All rates estimated for 21-44 year old male living in the Northeast. A ">" or "<" indicates a significant difference (p < .05) between white and black interviewers.