In Marcie L. Cynamon and Richard A. Kulka (Eds Health Survey Research Methods. Hyattsville, MD: Department of Health and Human Services, DHHS Publication No. (PHS) 01-1013,2001.

FEATURE PAPER

Comparing Telephone and Face-to-Face Interviewing in Terms of Data Quality: The 1982 National Election Studies Method Comparison Project

Melanie C. Green and Jon A. Krosnick

During the last three decades, American survey research has shifted from being dominated by face-to-face interviewing in respondents' homes (based on samples generated by block listing of residences) to telephone interviewing of samples generated by random digit dialing (RDD). Telephone interviewing has many practical advantages, including reduced cost, the possibility of a quicker turnaround time, and the possibility of greater standardization of administration through closer supervision of interviewers. Initially, telephone interviewing had another unique advantage as well: the possibility of computer-driven questionnaire presentation. With the advent of computer-assisted personal interviewing (CAPI), telephone interviewing's edge in this regard is gone, but it continues to maintain its other unique advantages.

Telephone interviewing also has obvious disadvantages. For example, showcards, which are often used to present response choices in face-to-face interviews, are more difficult to employ in telephone surveys, requiring advance contact and mailing of cards to respondents. Telemarketing has also made it more difficult to obtain response rates in telephone surveys as high as those obtained in face-to-face surveys. Furthermore, as of 1998, about 6% of the U.S. population did not have a working telephone in their household, prohibiting these individuals from participating in a telephone survey. Thus, it is not obvious that data quality in RDD telephone surveys will exceed that obtained from block-listed face-to-face surveys.

Over the years, a number of studies have been conducted to compare the quality of data obtained by these two modes. However, these studies have for the most part been atheoretical, looking for potential differences between modes with little conceptual guidance about what differences might be expected and why. Furthermore, the designs of these studies have often involved methodological confounds or limitations that restrict their internal validity and generalizability.

In this paper, we report the results of a new set of analyses exploring differences in data quality across modes. We begin by offering a series of theory-grounded hypotheses about possible mode differences, and we review what little evidence exists regarding their validity. We then report findings from an analysis of data from the 1982 National Election Study Method Comparison Project (MCP), an experiment designed to compare block-listed face-to-face interviewing with RDD telephone interviewing. Our focus is on three aspects of data quality: sample representativeness (gauged in terms of demographics), the amount of effort respondents devote to providing accurate answers (i.e., satisficing versus optimizing), and the extent to which people misportray themselves in socially desirable ways, rather than giving honest answers.

Hypotheses and Literature Review

Sample Quality

There are several reasons why sample representativeness may differ across modes. First, as we mentioned, telephone ownership is not universal, and people without telephones are automatically excluded from an RDD sample. These people may be disproportionately low in income (lack of money often prevents ownership of a working telephone), low in education (education is correlated with income), non-white (race is correlated with income), or young (young people usually have less disposable income and are more transient than older people).

In addition, the determinants of refusal to be interviewed may differ across modes. In particular, people who are socially disenfranchised and feel at greater social risk of manipulation and persecution may be reluctant to participate in telephone surveys, because it is more difficult to be sure exactly who is calling and what consequences might follow from the answers a respondent gives. But when an interviewer who seems to be friendly and trustworthy—and has documentation of his or her identification—appears on a person's doorstep, the importance and legitimacy of the enterprise may be more apparent, making such people less reluctant to participate. This may exacerbate the underrepresentation of individuals with lower social status in telephone surveys relative to face-to-face surveys.

Several studies have compared block-listing face-to-face surveys with RDD telephone surveys in terms of sample representativeness, three employing national samples (Groves & Kahn, 1979; Mulry-Liggan, 1983; Thornberry, 1987) and two employing local samples (Klecka & Tuchfarber, 1978; Weeks

The authors are at Ohio State University.

This research was commissioned by the National Election Study Board of Overseers. The authors wish to express their thanks to James Lepkowski and Robert Belli for their advice and to Kathy Cirksena for her support of this project. We are also grateful to Aldena Rogers and Chris Mesmer for their assistance with study design and data collection.

This research was supported by a grant from the National Science Foundation (SBR-9707741).

et al., 1983).[1] As expected, response rates were higher in the face-to-face surveys than in the telephone surveys. Whereas gender was unrelated to sampling method, RDD samples consistently overrepresented individuals aged 25 to 44 and underrepresented individuals age 65 or above, compared with block-listed samples. RDD samples consistently included a greater proportion of whites and a smaller proportion of nonwhites than block-listed samples, although this difference was not always statistically significant. Four studies found that RDD samples contained a greater proportion of high income respondents and a correspondingly smaller proportion of individuals with low incomes (Groves & Kahn, 1979; Klecka & Tuchfarber, 1978; Thornberry, 1987; Weeks et al., 1983). All five studies found that RDD samples contained more individuals with a great deal of formal education and fewer with little education. All of this suggests that RDD samples underrepresent segments of the population with lower social status. However, these studies only compared the two sampling methods with each other; none used benchmarks (e.g., census data) to assess which sampling method represented the population more accurately. The current study will provide such a comparison.

Satisficing

A second potential set of mode effects involves satisficing. Krosnick's (1991) theory of survey satisficing is based upon the assumption that optimal question answering involves a great deal of cognitive work (Tourangeau, 1984). Yet there are a variety of reasons why people may not expend all this effort. People can shortcut cognitive processing in one of two ways—via either weak satisficing or strong satisficing. Weak satisficing involves executing all the cognitive steps involved in optimizing, but less completely and with more bias. Strong satisficing involves seeking to offer responses that seem reasonable to an interviewer without having to do any memory search or integration of information at all.

The likelihood that a respondent will satisfice is thought to be a function of three classes of factors: respondent ability, respondent motivation, and task difficulty. People who have relatively limited abilities to carry out the cognitive processes required for optimizing and those who are minimally motivated to do so are the most likely to shortcut them. And people are most likely to shortcut when the cognitive effort demanded by a question is substantial.

Interview mode may influence the likelihood of satisficing by affecting respondent motivation and task difficulty. During a face-to-face interview, the interviewer's engagement in and enthusiasm for the process of exchange is likely to be conveyed through visual, nonverbal behavior and is likely to be infectious. Respondents whose motivation flags or who question the value of a survey can observe the interviewer's obvious seriousness and commitment to the enterprise, which may motivate them to generate thoughtful answers. Respondents interviewed by telephone cannot observe such nonverbal cues and so may be less motivated.

Telephone interviews are typically conducted at a quick pace, much quicker than face-to-face conversations normally go. The fast pace makes interpreting questions more difficult and may press respondents to generate answers more quickly and superficially compared with a slower pace of presentation. Therefore, telephone interviewing may increase the likelihood of respondent satisficing and may therefore decrease data quality.

Some previous research offers evidence that tests this hypothesis. Consistent with the satisficing hypotheses, some past studies found more acquiescence in telephone interviews than in face-to-face interviews (e.g., Groves & Kahn, 1979; Jordan, Marcus, & Reeder, 1980). Furthermore, various studies found that respondents said "don't know" significantly more often in telephone interviews than in face-to-face interviews (e.g., Aneshensel, Frerichs, Clark, & Yokopenic, 1982; Aquilino, 1992; Groves & Kahn, 1979; Herzog, Rogers, & Kulka, 1983; Jordan, Marcus, & Reeder, 1980; Locander & Burton, 1976), though one found no significant mode difference (Rogers, 1976). And a meta-analysis by de Leeuw (1992) confirmed a general trend toward fewer "don't know" responses in face-to-face interviews relative to telephone interviews.

Social Desirability

Another consideration relevant to mode differences is social desirability response bias–the notion that respondents sometimes intentionally lie to interviewers (Paulhus, 1984). There is reason to believe that social desirability response bias can vary depending upon data collection mode. Many past studies suggest that people are more likely to be honest when there is a greater distance (both physical and psychological) between themselves and their interviewers. Distance seems to be minimized when a respondent is being interviewed face-to-face in his or her own home. The more remote telephone interviewer has a lesser ability to convey favorable or unfavorable reactions to the respondent, and may therefore be seen as meriting less of the respondent's concern. Consequently, more social desirability bias might occur in face-to-face interviews than over the phone.

Surprisingly, however, the few studies done to date on mode differences do not offer support for this hypothesis. Some studies have found no reliable differences between face-to-face and telephone interviews in reporting of socially desirable attitudes (Aquilino, 1998; Colombotos, 1965; Rogers, 1976; Wiseman, 1972). Other work has found that reliable differences run opposite to the social distance hypothesis. For example, Aquilino (1994) found more reporting of socially undesirable behaviors in face-to-face interviews than in telephone interviews; Johnson, Haugland, and Clayton (1989) found similar results in a college student sample. And Groves (1979) found that respondents expressed more discomfort

[1]Gfroerer and Hughes (1991) compared RDD and block-listed samples. However, because different methods were used to oversample minorities in the two modes, this study does not provide an accurate test of demographic differences between modes. Similarly, Freeman, Kiecolt, Nicholls, and Shanks (1982) compared the two sampling methods but report demographics only for the head of the household rather than the respondent.

about discussing sensitive topics (e.g., racial attitudes, political opinions, and voting) over the telephone than face-to-face. This may occur because the telephone does not permit respondents and interviewers to develop as comfortable a rapport. Consequently, respondents may not feel they can trust their interviewers to protect their confidentiality as much as they might in face-to-face interviews, so they are more reluctant to reveal embarrassing facts. But the limited array of evidence on this point again calls for further testing.

Data

To test the hypotheses that sampling and data collection mode might affect sample representativeness, satisficing, and social desirability response bias, we analyzed data from the 1982 National Election Studies Method Comparison Project, a study designed to explore mode differences. Specifically, this study compared RDD telephone interviews to block listing sampled face-to-face interviews.

Data Collection

The 1982 MCP involved 998 complete or partial telephone interviews and 1,418 face-to-face interviews, all conducted during the three months following the 1982 congressional elections. All of the face-to-face interviews were conducted by the University of Michigan's Survey Research Center (SRC). The telephone interviews were randomly split between the Michigan SRC and the University of California at Berkeley's Program in Computer-Assisted Survey Methods. Essentially identical questionnaires were used for all interviews, although showcards used in the face-to-face interviews were replaced by spoken explanations in the telephone interviews. The survey was similar in length to other National Election Studies (which require approximately one hour to complete) and asked about a range of political beliefs, attitudes, and behaviors.

Measures

The 1982 MCP data set included measures that we could use to gauge three forms of strong satisficing (selection of no-opinion response options, nondifferentiation, and mental coin flipping) and social desirability response bias (Krosnick, 1991). To gauge the tendency to select a no-opinion response option, we employed the seven questions explicitly offering respondents such options. For each respondent, we calculated the percentage of these questions he or she was asked that were answered "don't know"/"haven't thought much," which was recoded to range from 0 to 1.

To gauge the tendency to nondifferentiate (i.e., rate a series of objects identically on a single rating scale), we focused on two batteries. The first was a set of seven 101-point feeling thermometers, and the second battery involved nine ratings of Ronald Reagan's personality traits. For the feeling thermometers, we recoded the 0–100 scale into 10 segments (0–10, 11–20, 21–30, 31–40, 41–50, 51–60, 61–70, 71–80, 81–90, 91–100). We then counted up the maximum number of identical ratings made by each respondent for each battery. These two numbers were then each rescaled to range from 0 to 1, and the two batteries' scores were averaged to yield a single assessment of nondifferentiation for each respondent.[2]

"Mental coin flipping" was assessed by examining the strength of association between presidential candidate preference (as gauged by the difference between attitudes toward Jimmy Carter and Ronald Reagan measured on feeling thermometers) and various predictors (all variables being coded to range from 0 to 1). Mental coin flipping should attenuate such associations.

Social Desirability Response Bias

Only five questions in the 1982 MCP seemed likely to have widely shared social desirability connotations, involving interest in politics, voting in previous elections, and support for government aid to blacks (the latter among Caucasians only). Admitting animosity toward African Americans is presumably not respectable among Caucasians. Additionally, interest and participation in politics are presumed to be civic virtues in this culture.

House Effects

We approached the assessment of mode differences in two ways. To gain the maximal statistical power by using the full array of cases, we compared the face-to-face interviews to the full set of telephone interviews. However, this comparison confounds mode with house, because Michigan conducted all the face-to-face interviews but half the telephone interviews were done by Berkeley. Therefore, although we did not expect significant house differences, if the standard interviewing practices at either institution differentially encouraged or discouraged satisficing, this comparison would be misleading about the effects of mode per se.

To deal with this problem, we conducted two additional sets of complementary analyses. First, we did statistical tests comparing the extent of satisficing in the Michigan and Berkeley telephone samples, to explicitly test for house effects. We also conducted less powerful tests comparing only the Michigan telephone respondents to the face-to-face respondents.

Results

Sample Comparability

We first examined whether the samples of respondents interviewed face-to-face and by telephone differed in their

[2]Although it may seem reasonable to use the variance of the responses as a measure of differentiation, we do not use this measure because we are interested in whether or not the respondent is giving different answers versus the same answers to a set of questions, not how extreme the differences in answers might be.

Table 1. Unstandardized regression coefficients predicting interview mode with demographic variables

Predictor	Unstandardized Regression Coefficients: All Predictors Entered Simultaneously	Each Predictor Entered Individually
Age	−.08*	−.06**
Education	.03	.07**
Gender	.04	.02
Race	−.10**	−.07**
Income	.16**	.12**
R^2	.02	
N	2097	

*$p < .05$
**$p < .01$

demographic characteristics. To do so, we initially regressed a dummy variable coded 0 for people interviewed face-to-face and 1 for people interviewed by telephone individually on one of the following demographic variables at a time: age (coded in years), education (coded in years), gender (coded 0 for males and 1 for females), race (coded 0 for Caucasians and 1 for others), and income (coded in a set of discrete categories representing ranges of dollars). As column 2 of Table 1 shows, the telephone respondents were significantly younger ($b = -.06$, $p < .01$), significantly better educated ($b = .07$, $p < .05$), significantly more likely to be Caucasian ($b = -.07$, $p < .01$), and of significantly higher income ($b = .12$, $p < .01$) than the face-to-face respondents. And in a regression using all the demographics to predict mode, the age, race, and income differences remained significant (see column 1 of Table 1).[3]

To assess which mode provided a more representative sample, we compared both samples to data from the 1980 U.S. Census. As shown in Table 2, the face-to-face sample matches the Census figures more closely than the telephone sample. For example, the face-to-face sample differs from the Census by 2.2% in terms of gender, whereas the telephone sample's discrepancy is 4.5%. Likewise, the face-to-face sample misses the percentage of Caucasians by .7%, whereas the telephone sample does so by 4.7%. And the average discrepancy between the face-to-face sample and the Census in terms of the percentages of people in the seven income categories is 3.2% on average, compared to 5.9% on average for the telephone sample. As expected, the telephone sample underrepresented low-income, less-educated, and non-white respondents more so than the face-to-face sample did.

The differences in samples created by block listing and RDD suggest that individuals in telephone samples should be more motivated and able to provide high-quality data. In order to provide a fair test of our hypotheses about data quality, we controlled for these variables in the analyses to follow,

[3]Throughout this study, significance tests of directional predictions are one-tailed, and tests of nondirectional predictions are two-tailed.

Table 2. Demographic characteristics of face-to-face and telephone samples and the nation (according to the 1980 U.S. Census)

Demographic	1982 MCP: Face-to-Face	1982 MCP: Telephone	1980 U.S. Census
Gender			
Male	44.7%	42.4%	48.6%
Female	55.3	57.6	51.4
Race			
White	88.5	92.5	83.1
Non-white	11.5	7.5	16.9
Age			
18–24	11.4	14.0	16.0
25–29	12.1	13.4	13.3
30–39	23.2	24.5	19.8
40–49	13.3	13.6	14.2
50–59	13.4	13.6	14.5
60–64	8.0	4.6	6.3
65 and over	18.6	16.3	15.9
Education			
Grade 8 or less	11.2	7.8	15.7
Grade 9–11	11.0	9.6	16.1
High school diploma	34.6	35.1	36.4
Some college	23.2	24.5	17.4
College graduate	20.0	23.0	14.4
Income			
Less than $5,000	10.7	4.4	13.2
5,000–9,999	13.7	12.1	15.9
10,000–14,999	14.8	11.0	15.3
15,000–19,999	10.9	12.0	14.1
20,000–24,999	13.7	13.8	12.4
25,000–34,999	17.3	22.4	15.7
35,000–49,999	11.2	14.2	8.6
50,000 and over	7.7	10.1	4.6

as well as other demographics that previous research suggests are related to the use of satisficing response strategies.

No-Opinion Responses

The first three columns of Table 3 display the mean proportions of no-opinion responses for the face-to-face respondents, the Michigan telephone respondents, and the Berkeley telephone respondents. The first row of the table reports results for the full sample and shows higher levels of no-opinion responding in the telephone samples (Michigan mean = 26%, Berkeley mean = 22%) than in the face-to-face sample (mean = 18%), consistent with the satisficing hypothesis. The difference is more pronounced when only the Michigan data are considered than when the Berkeley data are added in. There were in fact significantly fewer no-opinion responses in the Berkeley telephone data than in the Michigan telephone data ($b = .04$, $p < .05$, $N = 851$).

We tested the significance of the mode effect in two ways. In both cases, we conducted an OLS regression predicting the

Table 3. Analyses comparing rates of satisficing across modes in the 1982 NES

	Adjusted Means			Regression Coefficients							
		Telephone		Mode							
Sample	Face-to-Face	Mich.	Berk.	Mich.	Mich. & Berk.	Age	Educ.	Gender	Income	Race	R^2
No-opinion											
Full sample	.18	.26	.22	.09**	.07**	.03*	−.30**	.01+	−.07**	.05**	.18
				(1682)	(2095)						
Low education	.34	.54	.40	.21**	.15**	.03	−.46**	.07**	−.09+	.005	.09
				(350)	(410)						
High education	.14	.20	.18	.06**	.05**	.00	−.23**	.01	−.05**	.05**	.10
				(1332)	(1685)						
Nondifferentiation											
Full sample	.34	.37	.38	.02*	.03**	−.04**	−.02	−.01+	−.02+	−.01	.02
				(1684)	(2097)						
Low education	.34	.38	.40	.02	.05**	.01	.07	−.02	.06+	−.01	.03
				(351)	(411)						
High education	.35	.37	.39	.02*	.03**	−.05**	.00	−.01	−.03*	−.01	.02
				(1333)	(1686)						

N's appear in parentheses underneath coefficients. The effects of demographics are from an equation including the Michigan and Berkeley data.
+ $p < .10$
* $p < .05$
** $p < .01$

proportion of no-opinion responses with a dummy variable coded 0 for face-to-face respondents and 1 for telephone respondents, and controlling for a series of other demographic variables. This regression was done once only with the Michigan respondents and again with the Berkeley respondents folded in. The tests of mode effects are shown in the fourth and fifth columns of Table 3.

Consistent with the satisficing hypothesis, the mode effect was significant in both analyses, though a bit weaker when the Berkeley data were included ($b = .09$, $p < .01$ for the Michigan data only; $b = .07$, $p < .01$ including the Berkeley data). Furthermore, the effects of the demographic variables were largely consistent with prior research (Krosnick & Fabrigar, forthcoming). No-opinion responses were more common among respondents with less education ($b = -.30$, $p < .01$), those with lower incomes ($b = -.07$, $p < .01$), those who were older ($b = .03$, $p < .05$), those not Caucasian ($b = .05$, $p < .01$), and those who were female ($b = .01$, $p < .10$). These findings generally validate our analytic approach.

The satisficing hypothesis predicts that respondents' dispositions may interact with situational forces in determining the degree to which any given person will satisfice when answering any given question (Krosnick, 1991; Krosnick, Narayan, & Smith, 1996). That is, satisficing may be most likely when a person is disposed to do so *and* when circumstances encourage it. This logic suggests that the mode effect we observed might be strongest among respondents who were most disposed to satisfice. A great deal of research suggests that an especially powerful disposition in this regard is cognitive skills, which are very strongly correlated with years of formal education (Ceci, 1991) and can therefore be effectively measured in that way. We tested this interaction here.

Rows 2 and 3 of Table 3 display our findings when we tested this prediction by splitting the sample into respondents who had not graduated from high school and respondents with more education (for the rationale for this split, see Narayan & Krosnick, 1996). As expected, the mode effect was especially pronounced among the least-educated respondents. Looking only at the Michigan data, the average proportion of no-opinion responses increased from 34% in the face-to-face interviews to 54% on the telephone ($b = .21$, $p < .01$). The difference is smaller but nonetheless significant when the Berkeley data are folded in ($b = .15$, $p < .01$). The mode effect is much smaller in the highly educated subsample, though it is statistically significant there as well (Michigan data only: $b = .06$, $p < .01$; Michigan and Berkeley data: $b = .05$, $p < .01$).

Nondifferentiation

The second panel of results in Table 3 pertains to nondifferentiation. Here again, we see evidence consistent with the satisficing hypotheses. First, there was more nondifferentiation in the telephone samples (Michigan mean = .37, Berkeley mean = .38) than in the face-to-face sample (mean = .34). The latter rate is significantly lower than the telephone rate, whether we exclude the Berkeley data ($b = .02$, $p < .05$) or include it ($b = .03$, $p < .01$). The rate of nondifferentiation in the Michigan telephone sample was not significantly different from that in the Berkeley telephone sample ($b = .02$, n.s., $N =$

851), suggesting that all forms of satisficing were not uniformly less common in the latter.

Very little is known about the demographic correlates of nondifferentiation, other than the fact that it tends to be more common among less-educated respondents (Krosnick & Alwin, 1988; Krosnick, Narayan, & Smith, 1996; Rogers & Herzog, 1984). This trend was apparent here but was not statistically significant ($b = -.02$, $p > .10$); in fact, even the simple bivariate relation of education to nondifferentiation in the full sample was not significant ($b = -.01$, n.s., $N = 2403$). However, nondifferentiation was significantly or marginally significantly more common among respondents with lower incomes ($b = -.02$, $p < .10$), those who were younger ($b = -.04$, $p < .01$), and those who were male ($b = -.01$, $p < .10$).

When only the Michigan data were considered, the mode effect was no stronger in the least educated group ($b = .02$, n.s.) than in the more-educated group ($b = .02$, $p < .05$); (see the second and third rows of the bottom panel of Table 3). But when the Berkeley data were included, the mode effect was nearly twice as large in the least-educated group ($b = .05$, $p < .01$) as in the more educated group ($b = .03$, $p < .05$).

Mental Coin Flipping

To assess mental coin flipping, we gauged the strength of associations between variables via OLS regression. In all of them, the dependent variable was the difference between feeling thermometer ratings of Ronald Reagan and Jimmy Carter. Each regression included only a single predictor (in order to avoid problems that might be caused by multicollinearity and redundancy). The list of predictors included many well-documented strong correlates of candidate preference: party identification, job performance evaluations, perceptions of traits and emotions evoked, and assessments of personal finances and the national economy.

In the full sample, the average association was .44 for face-to-face respondents, compared with .41 for the Michigan telephone respondents and .43 for the Michigan and Berkeley telephone respondents combined. The difference between the face-to-face and telephone respondents was significant when only the Michigan telephone respondents are used ($z = 2.96$, $p = .002$) and when both the Michigan and Berkeley telephone respondents are used ($z = 1.90$, $p = .03$). Among the least-educated respondents, the mode effect was larger (as expected): the average association is .39 for the face-to-face respondents, .33 for the Michigan telephone respondents, and .36 for the Michigan and Berkeley telephone respondents ($z = 2.13$, $p = .02$, and $z = 0.90$, $p = .18$, respectively). Among more-educated respondents, the face-to-face mean association was .45, compared with .44 for both the Michigan only and the Michigan and Berkeley telephone respondents ($z = 1.36$, $p = .09$, and $z = 1.27$, $p = .10$, respectively).

Social Desirability

Two of the five tests of social desirability response bias differences by mode yielded significant results. Respondents interviewed by telephone reported higher interest in politics than respondents interviewed face-to-face ($b = .05$, $p = .01$). And Caucasians interviewed by telephone reported more support for government aid to blacks ($b = .06$, $p < .01$). Thus, telephone respondents were apparently more reluctant to report some socially undesirable interests and attitudes than were face-to-face respondents.[4]

Discussion

These analyses suggest that interview mode can affect both the sample representativeness and response patterns observed in surveys. In particular, individuals who were socially disadvantaged were undersampled in the telephone survey, relative to both the face-to-face survey and the population. Furthermore, data obtained from telephone interviews were more distorted by satisficing and by a need to appear socially desirable than were data obtained from face-to-face interviews. Individuals interviewed over the telephone showed more nondifferentiation and gave no-opinion responses more often, and these people showed an increased tendency toward socially desirable responding. These patterns are consistent with the notion that the rapport developed in face-to-face interviews inspires respondents to work harder at providing high-quality data, even when doing so means admitting something that may not be socially admirable. Given the concordance of these findings with ones from past studies, there seems to be a justifiable basis for confidence in their generalizability and validity. Furthermore, these findings suggest that there is validity to satisficing theory's claims and utility to its perspective for understanding survey phenomena.

The book is far from closed on the issue of interview mode and data quality, and the question remains an important one for survey researchers. Although telephone interviewing may be appealing to researchers because of the financial benefits, there may be significant costs associated with this method. Particularly when disenfranchising socially vulnerable members of a population is important (which it may always be), RDD telephone methods may not be worth the cost savings when budgets can permit block listing and face-to-face interviewing instead.

References

Aneshensel, C. S., Frerichs, R. R., Clark, V. A., & Yokopenic, P. A. (1982). Telephone versus in-person surveys of community health status. *American Journal of Public Health, 72,* 1017–1021.

Aquilino, W. S. (1992). Telephone versus face-to-face interviewing for household drug use surveys. *International Journal of the Addictions, 27,* 71–91.

[4]Parallel results were generated with only the Michigan respondents, except that the significant effect for interest in politics apparent in Table 5 falls to nonsignificance, likely due to the smaller sample size.

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

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology, 27,* 703–722.

Colombotos, J. (1965). The effects of personal vs. telephone interviews on socially acceptable responses. *Public Opinion Quarterly, 29,* 457–458.

De Leeuw, E. D. (1992). *Data quality in mail, telephone, and face to face surveys.* Amsterdam: TT-Publikaties.

Freeman, H. E., Kiecolt, K. J., Nicholls, W. L., & Shanks, J. M. (1982). Telephone sampling bias in surveying disability. *Public Opinion Quarterly, 46,* 392–407.

Gfroerer, J. C., & Hughes, A. L. (1991). The feasibility of collecting drug abuse data by telephone. *Public Health Reports, 106,* 384–393.

Groves, R. M. (1979). Actors and questions in telephone and personal interview surveys. *Public Opinion Quarterly, 43,* 190–205.

Groves, R. M., & Kahn, R. L. (1979). *Surveys by telephone: A national comparison with personal interviews.* New York: Academic Press.

Herzog, A. R., Rogers, W. L., & Kulka, R. A. (1983). Interviewing older adults: A comparison of telephone and face-to-face modalities. *Public Opinion Quarterly, 47,* 405–418.

Jordan, L. A., Marcus, A. C., & Reeder, L. G. (1980). Response styles in telephone and household interviewing: A field experiment. *Public Opinion Quarterly, 44,* 210–222.

Klecka, W. R., & Tuchfarber, A. J. (1978). Random digit dialing: A comparison to personal surveys. *Public Opinion Quarterly, 42,* 105–114.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5,* 213–236.

Krosnick, J. A., & Alwin, D. F. (1988). A test of the form-resistant correlation hypothesis: Ratings, rankings, and the measurement of values. *Public Opinion Quarterly, 52,* 526–538.

Krosnick, J. A., & Fabrigar, L. R. (forthcoming). *Designing good questionnaires: Insights from psychology.* New York: Oxford University Press.

Krosnick, J. A., Narayan, S., & Smith, W. R. (1996). Satisficing in surveys: Initial evidence. *New Directions for Evaluation, 70,* 29–44.

Locander, W., & Burton, J. P. (1976). The effect of question forms on gathering income data by telephone. *Journal of Marketing Research, 13,* 189–192.

Mulry-Liggan, M. H. (1983). A comparison of a random digit dialing survey and the Current Population Survey. In *Proceedings of the American Statistical Association, Section on Survey Research Methods* (pp. 214–219). Washington, DC: American Statistical Association.

Narayan, S., & Krosnick, J. A. (1996). Education moderates some response effects in attitude measurement. *Public Opinion Quarterly, 60,* 58–88.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598–609.

Rogers, T. F. (1976). Interviews by telephone and in person: Quality of responses and field performance. *Public Opinion Quarterly, 40,* 51–65.

Rogers, W. L., & Herzog, A. R. (1984). Response style characteristics and their relationship to age and item covariances. Unpublished manuscript, Institute for Social Research, Ann Arbor, MI.

Thornberry, O. T. (1987). *An experimental comparison of telephone and personal health interview surveys.* Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics.

Tourangeau, R. (1984). Cognitive sciences and survey methods. In T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey methodology: Building a bridge between disciplines.* Washington, DC: National Academy Press.

Weeks, M. F., Kulka, R. A., Lessler, J. T., & Whitmore, R. W. (1983). Personal versus telephone surveys for collecting household health data at the local level. *American Journal of Public Health, 73,* 1389–1394.