Interviewer Effects in Telephone Surveys

CLYDE TUCKER

EVEN the best interviewer training will not eliminate all differences in the behaviors of interviewers. To the extent that they persist, these differences can produce variation in the results of surveys from interviewer to interviewer. This variation, usually labeled "interviewer effects," can increase the amount of error associated with the measurement of variables in a survey. Studies which focus on between-interviewer variation assess the magnitude of these effects and the variables for which the effects are likely to be the greatest. This study examines the between-interviewer variation existing in the national telephone surveys conducted by CBS News and *The New York Times* in 1980, and unlike most of the previous studies of this type, it analyzes the same variables across several polls.

Several consistent findings have emerged from the previous studies of interviewer effects. The first is that interviewer effects do indeed exist. When these effects are present, the total variance on items can be substantially larger than the variance due to sampling alone (Han-

Abstract The interviewer effects for selected items from a number of national polls conducted by CBS News and *The New York Times* in 1980 were examined. These effects were found to be generally quite small and somewhat inconsistent from poll to poll. The inconsistencies were explained by variable associations with the nonrandom regional distribution of respondents and the political context in which the measurements were obtained. There was some evidence of respondent-interviewer interactions for certain items.

Clyde Tucker is Assistant to the Director of the CBS News Election and Survey Unit. An earlier draft of this article was presented at the Annual Conference of the American Association for Public Opinion Research, Buck Hill Falls, Pennsylvania, May 28–31, 1981. The author wishes to express his appreciation for advice given on earlier drafts by Warren Mitofsky, Kathleen Frankovic, Robert Groves, Murray Edelman, Mohammed Yusuf, and two anonymous reviewers. Thanks also go to Solomon Barr, Wayne Reedy and Carolyn Stroock for their assistance in the statistical analysis and preparation of the manuscript.

Public Opinion Quarterly Vol. 47:84–95 © 1983 by the Trustees of Columbia University
Published by Elsevier Science Publishing Co., Inc. 0033-362X/83/0047-84/$2.50

son and Marks, 1950; Stock and Hochstim, 1951; Kish, 1962; Sudman and Bradburn, 1974; Bailey et al., 1978; Groves and Kahn, 1979; Groves and Magilavy, 1980).

Not only has the existence of interviewer effects been established, but it has also been demonstrated that the magnitude of these effects differs from item to item. As Groves and Magilavy point out, originally it was thought the difference turned on whether the question was factual or nonfactual. They suggest, however, that the effects are actually a function of the amount of interviewer interference which is possible. This interference can take a variety of forms and can be a factor in both factual and nonfactual items. Questions which seem to be most susceptible to interviewer interference are those which concern sensitive topics leading to resistance in asking or responding, open-ended questions (especially those involving probes), and questions requiring a rating or subjective assessment from the interviewer.

A number of studies have also found that interviewer effects are related to characteristics of the interviewer and respondent. Older respondents are most open to interviewer effects (Hanson and Marks; Groves and Magilavy). Interviewer effects seem to be related to interviewer competence (determined in a variety of ways) and an interviewer's prior expectations of survey results. There is also evidence to suggest that younger interviewers are less susceptible to interviewer effects. Finally, several studies have shown that interviewer effects can be the product of an interaction between interviewer, respondent, and item characteristics (Athey, et al., 1960; Williams, 1964; Dohrenwend, et al., 1968–69; Schuman and Converse, 1971; Hatchett and Schuman, 1975–76; Freeman and Butler, 1976; Schaffer, 1980; Campbell, 1981).

Perhaps the most interesting aspect of these studies is the methodological issues they raise. Estimating interviewer effects has turned out to be a very complicated task. A large part of the problem is design limitations. Most surveys, whether they are done by telephone or personal interview, are constrained by time and cost. Furthermore, they are not usually conducted for methodological purposes so that ease of estimation of interviewer effects is not considered a high priority, and few are willing to jeopardize the quality of a survey by altering procedures in order to measure what may be relatively small effects. There are also the practical problems associated with any survey in a real-life situation which hinder the measurement of interviewer effects.

In addition to the problems imposed by design limitations, there are the assumptions of the statistical procedures used in estimating interviewer effects. Often the assumption of equal variances across inter-

viewers will be violated, especially when proportions are involved. Sometimes this is a result of small numbers of respondents per interviewer or a failure in the randomization process, but it can also reflect true interviewer differences. Other assumptions which must be considered include uncorrelated errors between interviewers and the absence of a correlation between response and sampling deviations (Groves and Magilavy).[1]

Design

This study relies on data from 11 national CBS News/*New York Times* polls taken during the course of the 1980 election year.[2] These polls were conducted over a three- to six-day period with telephone numbers generated by a two-stage clustered sampling technique.[3] The sample sizes for the polls ranged from 1100 to almost 3100 in the fall pre-election poll. These interviews were gathered by 40 to 60 interviewers each day.

In all, 300 interviewers were involved in the 11 polls.[4] Some worked on a number of polls while others were present for only one. Each interviewer was assigned a unique identification number, and

[1] When investigating interaction between interviewer and respondent characteristics by item, one needs to be especially careful to see that all cells have enough cases in the subclasses of the relevant variables. These interactions can produce correlated errors between interviewers unless appropriate techniques are used to isolate their effects.

[2] The polls used for this study were conducted on the following dates (the reduced registered and total sample sizes and number of interviewers are also given):

Poll Dates	*Number of Interviewers*	*N's*
Jan. 9–13, 1980	35	999
Feb. 13–17, 1980	36	1092
Mar. 12–15, 1980	37	872
Apr. 10–14, 1980	41	1281
June 18–22, 1980	41	1171
Aug. 2–7, 1980	44	1386
Sep. 10–14, 1980 (Sep. poll)	43	1603 (reg. and unreg.) 1098 (reg.)
Sep. 19–21, 1980 (predebate one)	43	912 (reg. and unreg.) 633 (reg.)
Sep. 22–24, 1980 (postdebate one)	40	891 (reg. and unreg.) 643 (reg.)
Oct. 16–20, 1980 (Oct. poll)	58	1635 (reg. and unreg.) 1175 (reg.)
Oct. 30–Nov. 1, 1980 (preelection poll)	61	2309 (reg. and unreg.) 1703 (reg.)

[3] See Mitofsky (1980) and Waksberg (1978).

[4] While 300 interviewers were used in these polls, only 245 completed interviewer questionnaires. Only data from the 245 are analyzed here to provide continuity for later studies.

this number was recorded on every interview he or she completed. The training session lasted two to three hours, after which the new interviewers were allowed to begin interviewing. During the first few interviews, they were monitored by experienced interviewers. Any problems discovered at this time were discussed with the interviewer. Further monitoring was then done to ensure that correct procedures were being followed.

On the first night of a poll, both new and experienced interviewers received a briefing. Each of the questions was read aloud, and any unusual aspects of a question were pointed out. Most of the items had already been pretested so that inadequate or confusing questions had been eliminated. The skip patterns in the questionnaire were also discussed during the briefing.

An effort was made to distribute phone numbers in a random fashion, given the limits existing in the survey situation. Some phones could not reach every area of the country, so that an interviewer, on a given night, might not get to call every part of the country. Over the course of the poll, however, interviewers should have obtained a number of interviews from each region.

Sheets with telephone numbers were randomly ordered within region. Poll managers supervised the random distribution of these sheets. After the first call on numbers, those that had been "no answers" were redistributed randomly. This procedure was followed until four calls had been made on all numbers. Whenever possible, an appointment call was made by the interviewer who had set up the appointment.

In several of the polls, refusals were re-called. Approximately 2 to 3 percent of the respondents in these polls had originally been refusals. An examination of two surveys which contained refusals shows that there are differences between former refusals and the rest of the sample. Former refusals tend to be older and less educated than other respondents. They are also more likely to give "don't know" responses to substantive questions.

While the refusals are spread over a number of interviewers, a few of the interviewers do have a significant portion of their interviews coming from this group (never more than 30 percent). Results from analyses both including and excluding interviewers with more than 10 percent of their respondents from the former refusals show few differences. In other words, interviewers with a large number of former refusals as respondents do not significantly add to the interviewer effect. This knowledge, coupled with the fact that former refusals are such a small portion of the sample, permits us to be relatively unconcerned over their inclusion in the analysis.

The analysis of interviewer effects depends upon a model of the variance of a survey response given by Hansen et al. (1961). Perhaps, O'Muircheartaigh (1977) offers the clearest exposition of this model. Essentially, the problem is one of estimating the intra-interviewer correlation between response deviations, ρ_{int}. In terms of an analysis of variance model when interviewer is the sole independent variable, $(B\text{-}W)/n$ (between mean square-within mean square/no. of respondents) is a good estimator of $p_{int}\,\sigma_d^2$ (where σ_d^2 is the variance of the response deviations) if the survey is interpenetrated, ρ'_{int} (the correlation of response deviations between interviewers) is near zero and the correlation between sampling and response deviations is also minimal (Fellegi, 1964; Groves and Magilavy, 1980). The surveys considered in this study are interpenetrated. For the purpose of this analysis, the two assumptions are accepted while it is recognized that they may not be completely valid.

While $(B\text{-}W)/n$ controls for interviewer load, it is still dependent upon the variance in the variable being measured. Instead, a measure developed by Leslie Kish (1962), which controls for this factor, is employed. The form of this estimator is given by Groves and Magilavy as,

$$\frac{(B\text{-}W)/n}{((k\text{-}1)\,B\ +(n\text{-}1)\,kW)/(nk\text{-}1)}$$

where k is the number of interviewers and n is the average number of respondents per interviewer.[5] This estimator has an expected value of $\frac{\rho_{int}\,\sigma_d^2}{\sigma_d^2\ +\ \sigma_s^2}$ (where σ_s^2 is sampling variance) when the two assumptions mentioned earlier are met. It is labeled p^*_{int} and is used as the principal measure of interviewer effects in this paper.

Randomization and Comparable Distributions

As mentioned earlier, a first requirement for an accurate estimate of interviewer effects is that each interviewer have a random subsample

[5] The components of the mean squares, between and within, are given by Kish (1962). Between is $s_s^2+ks_a^2$, and within is s_s^2. In this case, s_s^2 is the sampling variance, s_a^2 is the response variance due to interviewer effects, and k is the number of interviews per interviewer. Ordinarily the following formula would be used to give an estimate of n which takes into account the varying cell sizes:

$$n^* = \frac{(\Sigma n_i)^2 - \Sigma n_i^2}{(\Sigma n_i)(k\text{-}1)}$$

But this computation becomes tedious, so an average correction was applied to the average cell size for each item. This is the same procedure used by Kish (1962). The correction applied was subtracting .072 from the average. See also Felligi (1964) for estimates of variance components.

of respondents. The method for distributing phone numbers was designed to accomplish this randomization as much as possible within the constraints of the survey situation. This procedure definitely broke down in the case of the West. Western numbers, from those states in the Western census region, were not distributed before it was 6 P.M. in that area (9 P.M. in the East). By 11 P.M. in New York, the only numbers being called were in the West. Some of the interviewers left at that time while others remained until 1 A.M. Consequently, one group of interviewers had a smaller proportion of Western interviews than the rest. Since there is a difference in response patterns from Western respondents compared to those from other regions, all of these interviews were excluded from the analysis.[6]

The response patterns for four demographic variables were examined as a test for the random distribution of respondents across interviewers. These variables are sex, region, size of place, and number of adults in the household.[7] The variable which shows the greatest variance from interviewer to interviewer is region. This leads us to suspect that randomization by region was not successfully accomplished. How important this failure is depends on the level of the associations between region and other variables. Because of this problem, separate p^*_{int} values for respondents in the South and non-South will be given along with p^*_{int} values for the total sample.

One possible explanation of the nonrandom distribution of interviews is in the refusals which each interviewer encounters. Ideally, these refusals will be randomly distributed among the interviewers, but this cannot be assumed. Just as differences in interviewer behavior produce variation in the results from completed interviews, they can lead to nonrandom refusal patterns. In fact, the distribution of refusals, itself, can be analyzed for evidence of interviewer effects. However, in this case, the distribution of refusals was examined only with the hope of explaining the nonrandom distribution of completed interviews.

To consider the possible differences in refusals for interviewers, the refusals from two polls (January and April) were analyzed. Again, the distributions on the four demographic variables were examined.[8] Only

[6] Other respondents were also eliminated. Besides the respondents interviewed by those not completing the background questionnaire, interviews from interviewers who had fewer than 10 completes in the total sample were excluded. Respondents who were a part of a Republican oversample in January, February, and March were also eliminated.

[7] Region was dichotomized into South and non-South (Northeast and Midwest). Size of place was dichotomized into rural and nonrural (small cities, suburbs, and large cities). Number of adults in the household was not recoded.

[8] *Interviewers with five or more refusals were included here.*

in the case of region were there any differences, but the distribution of refusals for particular interviewers was almost identical to their completes. While this leads us to further question the randomization procedures, it does indicate that, given the way numbers were distributed, refusals were randomly distributed.[9]

Interviewer Effects

Using the analysis of variance procedure which ultimately yields the p^*_{int} measure, an examination of the interviewer effects for a number of variables across the 11 polls was undertaken. These variables were chosen according to two criteria—substantive importance (those questions found to be important in the context of the 1980 campaign) and frequency of use. Only questions appearing in at least five polls were included in this analysis.

Besides the overall p^*_{int}, two other p^*_{int}s are calculated—one for respondents in the South and another for the non-South respondents (those in the East and Midwest). This regional distinction was chosen because attitudes, behavior, and demographics have often been shown to be different for the South compared to the rest of the country.[10] The problem is that cell sizes can be quite small, especially in the South subclass. Occasionally, there will only be one case in a cell. This increases the unequal cell variance problem and results in some unstable estimates of p^*_{int}.

The ρ_{int}s for the items in this study are shown in Table 1. A cursory examination shows that, in most cases, interviewer effects are not significant. This is true whether one considers the p^*_{int}s for the total or for the regional subclasses. No more than a third of the items on a single poll have a significant p^*_{int}, and the items which are significant vary from one poll to another.[11] The fact that most of the items from September through the preelection poll were asked only of registered respondents does not appear to be important here.

Inconsistencies in the total and non-South estimates for an item

[9] Another possible reason for the failure of randomization is the varying "no answer" patterns for the interviewers, but this is unlikely. Only if interviewers lied about making calls or did not let the phone ring an equal number of times (10) would differences result. Furthermore, "no answers" are randomly distributed up to three more times so that any interviewer idiosyncracies should be eliminated.

[10] See Campbell et al. (1960), Lane (1959), and Elazar (1972).

[11] One reviewer felt I should mention that the measures of interviewer effects are related because the items are not independent replications. Of course, this is always the case when only one sample is involved. The measures are, however, independent across polls.

across polls can partially be attributed to variable associations with region. Item 10, the proportion paying a lot of attention to the campaign, is a good example. The significant p^*_{int} for this item in September seems to be a function of its association with region. The same is true in the case of the significant p^*_{int} for item 3, also in September.

In a number of instances, however, regional associations do not explain the results. Item 1, for example, is always associated with region, yet the estimates vary. No regional associations exist for Item 2, but, again, the estimates differ. For other items, there are cases when a significant p^*_{int} occurs in the absence of a regional association but not when a regional association exists.

Differences between estimates for some items may be attributed to context. For example, respondents may have been more susceptible to interviewer influence when the registration question (item 3) was asked in January, than they were once the campaign was in full swing. Interviewer effects for item 4, the proportion approving Carter's performance, are highest in June, shortly after his loss of several primaries. This is also the case for item 16, the proportion expressing a preference for Carter. In item 7, opinion of Anderson, effects are greatest in February, when he was just beginning to receive exposure, after his debate with Reagan in September, and just before the election (and after the Reagan/Carter debate).

The South p^*_{int}s are not affected by regional associations, but they still display a pattern similar to the others. The only difference is that they are more extreme in either direction than those for the total and non-South. This is probably a result of skewed proportions in cells with a small number of cases. Yet, even here, few significant values can be found. When the South p^*_{int} is similar to the other two estimates, it is an indication that region is not a factor in the estimates.

A significant total p^*_{int} does not always coincide with significance in the subsamples. This situation exists when regional distributions differ from interviewer to interviewer producing large effects. These effects disappear once region is controlled. On the other hand, when a significant total p^*_{int} and a subsample p^*_{int} occur at the same time, it usually indicates that the interviewer effects in the subsample are the source of significant effects in the total. A significant p^*_{int} in one subsample but not in the total means the subsample effects are too weak to appear in the total sample. If both subsamples have large p^*_{int}s but the total does not, then the subsample effects could be canceling each other in the total sample where the distributions around the cell means have changed.

The total p^*_{int} does not always fall between the values for the two subsamples. As indicated above, the total p^*_{int} will be larger than

Table 1. Intra-interviewer Correlations (ρ^*_{iw}) for Selected Categories of Items from 1980 CBS News/New York Times National Polls

Question	*Jan 80*[a]	*Feb 80*	*Mar 80*	*Apr 80*	*Jun 80*	*Aug 80*	*Sep 80*[b]	*Pre-debate one*	*Post-debate one*	*Oct 80*	*Pre-election 1980*
1. Proportion identifying as Democrats											
Total	.005	001	.012	.011	.004	.004	−.009	−.008	−.020	−.007	.003
Non-South	−.008	.004	.013	025*	.021	.001	−.017	−.008	−.033	−.006	−.012
South	−.006	−.017	.017	−.008	−.035	.009	.001	−.012	−.039	−.024	.015
2. Proportion with family income between $15–25,000											
Total	−.001	.014*	.006	−.004	.012	.001	.007	.024	−.014	−.014	
Non-South	−.013	.009	.004	−.010	.011	−.007	.011	.022	−.035	−.015	
South	.017	.044*	−.015	−.002	.052*	.008	.025	.054	.019	.011	
3. Proportion saying name recorded in the voter registration book where they live											
Total	.018*	−.006	−.007	−.007	−.010	−.006	.013*	.009	.015	−.001	.006
Non-South	.008	.014	−.004	.004	.004	.003	.040*	−.007	.001	−.005	.001
South	−.005	−.016	−.001	−.002	−.031	−.016	−.001	−.010	.022	−.001	−.002
4. Proportion approving Carter's performance as President											
Total	−.004	.010	.005	.014*	.026*	.009	.006				
Non-South	.014	.023	.004	.023*	.014	.009	.016				
South	−.003	.008	.010	.008	.020	.030	−.015				
5. Proportion having a favorable opinion of Carter											
Total	.017*	.002	.008	−.004	.026*	.007	.007	.011	−.002	.000	.001
Non-South	.032*	−.011	−.001	.005	.019	.010	.020	−.007	−.007	−.009	−.011
South	.006	.043*	.025	−.029	− 005	.026	−.015	.088*	−.005	.018	−.001
6. Proportion having a favorable opinion of Ronald Reagan											
Total	.004	.003	.029*	.003	−.011	−.003	−.005	.021	.009	.006	.005
Non-South	.000	.014	.028	−.003	−.017	.000	−.012	.030	.025	.029*	−.001
South	.011	−.006	.006	.012	−.034	−.011	.000	.007	.023	.024	−.013
7. Proportion having a favorable opinion of Anderson											
Total	−.004	.017*	−.012	−.007	−.003	−.004	.005	.016	.011	.004	.020*
Non-South	−.005	.012	.004	−.017	−.001	−.010	.008	.005	.014	−.032	.011
South	.033	.040*	−.005	−.023	.001	.039*	.007	.006	.084*	.008	.047*
8. Proportion naming inflation, unemployment, recession, or some other economic problem as the most important problem facing the country[c]											
Total	−.002	−.001	.022*	.004		−.001	.003	.005		−.003	
Non-South	−.014	−.010	.003	.015		−.002	.015	.029		.003	
South	− 008	−.008	.072*	−.026		−.014	−.038	.045		−.017	
9. Proportion approving of Carter's handling of the Iranian situation											
Total	−.002	.003	.016	−.003	.010	.014*				.028*	.005
Non-South	.014	.020	.012	−.018	.012	.018*				.018	.003
South	−.018	−.008	.015	.004	−.015	.015				−.005	−.006

10. Proportion paying a lot of attention to the Presidential campaign											
Total	−.008	−.005	−.007	.012	−.002	.005	.016*				−.002
South	−.001	−.001	.012	.006	.008	−.008	.002				.000
South	−.022	.000	−.026	.009	.014	.006	.025				−.011
11. Proportion saying inflation a more important problem than unemployment											
Total				.001	−.003	.017*	−.010			.009	
Non-South				.006	.014	.016	−.015			.012	
South				−.001	−.007	.004	−.020			.013	
12. Proportion favoring the Equal Rights Amendment											
Total		−.003	.024		−.002	.028*	−.006	.007	.009	.007	
Non-South		.009	042*		−.006	.036*	−.018	.038	.039	.019	
South		−.007	−.005		−.020	.004	.005	−.044	−.011	−.032	
13. Proportion with family financial situation worse today than a year ago											
Total		.014*	.002	.019	−.004	.003	−.003	−.001		.010	
Non-South		021	.013	.010	−.017	.000	−.004	.005		−.018	
South		.014	.009	.034*	−.028	−.004	−.015	−.012		.063*	
14. Proportion who voted in 1976 presidential election											
Total				.012	.011	−.004	−.004	.013	−.009	−.017	.010
Non-South				.021*	.010	.006	−.024	−.026	−.035	−.009	.014
South				−.008	−.016	−.004	.033	−.026	−.008	−.021	.020
15. Proportion with a union member in the household											
Total					.004	.016	−.001[f]			.009	.002
Non-South					.006	.019	−.012			.006	−.006
South					.037[f]	−.018	−.002			.032	.027
16. Proportion who would vote for Carter if the presidential election were held today											
Total					.016*	.015*	.008	.000		.000[f]	.001[f]
Non-South					.034*	.020*	.038*	.020		−.016	.007
South					−.011	.009	−.010	−.001		.013	.013
17. Proportion who will definitely vote for president											
Total						−.002	.006	−.005	−.019	.002	.009
Non-South						.006	.016	.002	−.020	.013	.022*
South						−.001	.038	.022	−.006	−.027	−.019
18. Proportion believing it is possible to balance the federal budget and cut income taxes at the same time											
Total						.002	.002	.004	−.010	−.011	
Non-South						.007	.021	−.015	.004	−.009	
South						.000	.021	−.019	−.039	−.031	

* $p < .05$.

[a] The items in the polls from January through August were asked of all respondents.

[b] Most items in the polls from September through the preelection poll were asked only of respondents who were registered to vote. If an item was asked of all respondents, it is indicated by *t*.

[c] Originally open-ended given close-ended codes.

those for the subsamples when the subsample distributions differ between interviewers. The total p^*_{int} is smaller than those for the subsamples when their effects cancel out in the total sample. These results are to be expected when ratios of variances are involved. They cannot simply be averaged like means.

Discussion

This study has assessed the magnitude of interviewer effects for a number of items across 11 national polls conducted by CBS News and *The New York Times* in 1980. Some effects were found to exist, but they were generally quite small. Even with the unequal regional distribution of respondents among the interviewers, the effects were comparable to those from previous studies of telephone surveys. Unlike most of these other studies, interviewer effects for an item were examined over several polls. There was some inconsistency in effects from poll to poll. In some cases, effects were related to the political context in which the measurements were obtained. In others, the variable associations with region explain the inconsistencies.

Effects were small, but as Groves and Magilavy point out, since interviewer load is large in telephone surveys, the cumulative effects of intra-interviewer correlations can be substantial. We need to be particularly mindful of the fact that interviewer effects may be tied to context. Since, in some cases, interviewer effects were restricted to one or the other subsample, the possibility of interactions between specific respondent and interviewer characteristics must be given greater consideration. Finally, the failure to achieve complete randomization with respect to the distribution of telephone numbers points to the need for greater control over design procedures so that interviewer effects can be more accurately measured.

References

Athey, K. R., J. E. Coleman, A. P. Reitman, and J. Tang
1960 "Two experiments showing the effect of the interviewer's racial background on responses to questionnaires concerning racial issues." Journal of Applied Psychology 44:244–46.

Bailey, L., T. F. Moore, and B. Bailar
1978 "An interview variance study for the eight impact cities of the national crime survey's cities sample." Journal of the American Statistical Association 73:16–23.

Campbell, A., P. E. Converse, W. S. Miller, and D. E. Stokes
1960 The American Voter. New York: Wiley.

Campbell, B.
1981 "Race-of-interviewer effects among Southern adolescents" Public Opinion Quarterly 45:231–44.

Dohrenwend, B. S., J. Colombotos, and J. Dohrenwend.
1968–1969 Social distance and interviewer effects." Public Opinion Quarterly 32:410–22.

Elazar, D. J.
1972 American Federalism: A View from the States. New York: Crowell.

Fellegi, I. P.
1974 "Response variance and its estimation." Journal of the American Statistical Association 69:496–501.

Freeman, J., and E. W. Butler
1976 "Some sources of interviewer variance in surveys." Public Opinion Quarterly 40:79–91.

Groves, R. M., and R. L. Kahn
1979 Surveys by Telephone. New York: Academic Press.

Groves, R. M., and L. J. Magilavy
1980 "Estimates of interviewer variance in telephone surveys." Proceedings of Survey Research Methods Section, American Statistical Association, 622–27.

Hansen, M. H., W. N. Hurwitz, and M. A. Bershad
1961 "Measurement errors in census and surveys." Bulletin of the ISI 38:351–74.

Hanson, R. H., and E. S. Marks
1958 "Influence of the interviewer on the accuracy of survey results." Journal of the American Statistical Association 53:635–55.

Hatchett, S., and H. Schuman
1975–1976 "White respondents and race of interview effects." Public Opinion Quarterly 39:523–28.

Kish, L.
1962 "Studies of interviewer variance for attitudinal variables." Journal of the American Statistical Association 57:92–115.

Lane, R. E.
1959 Political Life. New York: The Free Press.

Mitofsky, W. J.
1980 "Methods for telephone surveys." Unpublished CBS Memorandum.

O'Muircheartaigh, C. A.
1979 "Response effects." Pp. 193–239 in C. A. O'Muircheartaigh and C. Payne (eds.), The Analysis of Survey Data: Exploring Data Structure. New York: Wiley.

SAS Institute
1979 SAS User's Guide. Raleigh, N.C.: SAS Institute, Inc.

Schaffer, N.
1980 "Evaluating race-of-interviewer effects in a national survey." Sociological Methods and Research 8:400–19.

Schuman, H., and J. Converse
1971 "The effects of black and white interviewers on black respondents in 1968." Public Opinion Quarterly 35:44–68.

Stock, J. S., and J. R. Hochstim
1951 "A method of measuring interviewer variability." Public Opinion Quarterly 15:322–34.

Sudman, S., and N. H. Bradburn
1974 Response Effects in Surveys: A Review and Synthesis. Chicago: Aldine.

Waksberg, J.
1978 "Sampling methods for random-digit dialing." Journal of the American Statistical Association 73:40–46.

Williams, J. A., Jr.
1964 "Interviewer-respondent interaction: a study of bias in the information interview." Sociometry 27:338–52.