DOCUMENT RESUME

ED 241 596 TM 840 148

AUTHOR Rodgers, Willard; Herzog, Regula
TITLE Interviewer Effects on Responses of Younger and Older Respondents.
PUB DATE 20 Nov 83
NOTE 19p.; Paper presented at the Annual Meeting of the Gerontological Society of America (San Francisco, CA, November 20, 1983).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adults; *Age Differences; Interviews; *Older Adults; Reliability; Research Methodology; *Research Problems; *Surveys
IDENTIFIERS *Interviewer Effects; *Respondent Behavior

ABSTRACT

Using data collected through telephone interviews with a national sample of adults, this study searched for evidence as to whether interviewers have stronger effects on the responses given to a wide range of questions by older people than on the responses of younger people. Responses to 30 items for which significant interviewer effects had earlier been reported were examined. Significant age-by-interviewer interactions were detected for eight of the 30 items. The set of eight items included both interviewer observations and the direct responses of interviewers, and, among the latter, responses to both open- and closed-ended questions. Older and younger respondents were examined separately to determine the nature of the age-by-interviewer interactions. No pattern was found for the interviewer observations, but for the direct responses to questions, the median interviewer effect was more than twice as strong for older respondents as for younger respondents. Based on all the evidence from this study, it appears that older respondents may be somewhat more susceptible to interviewer effects than younger respondents. (BW)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED241596

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it
Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy

# Interviewer Effects on Responses of Younger and Older Respondents

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
W. Rodgers
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

by
Willard Rodgers and Regula Herzog

Survey Research Center
Institute for Social Research
The University of Michigan

Paper presented as part of symposium on
"Survey Research on Older Populations."
36th Annual Scientific Meeting of the Gerontological Society of America,
November 20, 1983, San Francisco.

TM 830 148

It has often been assumed that differences in interviewers' behaviors and attitudes may be an important factor affecting responses to survey questions; and much interviewer training is aimed at standardizing interviewer behavior in order to minimize biasing influence. Survey methodologists have been studying interviewer effects for several decades (e.g., Freeman and Butler, 1976), and the Bureau of the Census has an equally long history of measuring the impact of interviewer variability on census statistics (Bailar, 1976; Hanson and Marks, 1958) and, more recently, in sample surveys conducted by the Bureau (Bailar, Bailey, and Stevens, 1977; Bailey, Moore, and Bailar, 1978). Similarly, in market research it has long been accepted that variation in response to survey questions caused by differential interviewer questioning techniques or interpretation of responses can be a major source of bias or nonsampling error (Boyd and Westphal, 1955, 1965, 1970; McKenzie, 1977). Considerable research is also carried out by survey methodologists demonstrating that individual differences of interviewers as well as interactions of respondent and interviewer characteristics produce different responses from the same respondents (e.g., Cahalan, et al., 1947; Dohrenwend, et al., 1968; Freeman and Butler, 1976; Groves and Kahn, 1979; Hyman, 1954; Kish, 1962; Stock and Hochstim, 1951; Sudman and Bradburn, 1974; Sudman, et al., 1977). Although these studies differ substantially in populations sampled, interviewing staffs, types of variables analyzed, and in the particular statistics they examined, in combination they suggest that as much as 5 to 10 percent (or even more) of the variation in respondent reports of certain attitudes, behaviors, and characteristics may be attributed to interviewer differences.

Thus, variation in responses due to differences among interviewers may sometimes amount to a significant and important source of error in survey research. Moreover, such effects are likely to be particularly problematic if they tend to be correlated with certain basic characteristics of the respondent, such as age, sex, or education. With regard to age, there is indeed some research evidence which suggests that older people may be more susceptible to interviewer effects than younger people. For example, in various experimental situations older subjects are generally found to conform more readily to social influence manipulations and are more eager to please (Herzog, 1979; Klein, 1972; Klein and Birren, 1972). In addition, they traditionally tend to score higher on measures of social desirability (Herzog and Rodgers, 1981; Campbell et al., 1976). It is quite possible that each of these basic tendencies could result in increased susceptibility to influence by a survey interviewer.

Despite the plausibility of this hypothesis, however, few studies of interviewer variance have even considered possible age differences in respondent susceptibility to interviewer effects, and those which have done so provide only indirect or limited evidence bearing on the issue. For example, Sudman and Bradburn's (1974) comprehensive review of data on interactions between respondent and interviewer characteristics on response effects revealed no evidence of age effects for adult respondents, perhaps because the data available (e.g., Benney, et al., 1956; Ehrlich and Riesman, 1961; Hanson and Marks, 1958) were too limited in age range to either prove or disprove an interaction effect due to respondent and interviewer age. Similarly, in two recent and more direct studies of interviewer variability, Bailey and his colleagues (1978) examined the effects of National Crime Survey interviewers on survey statistics by sex and race of respondents, but not by age; and Sudman et al. (1977), in spite of giving passing recognition to the possible

influence of age on interviewer contribution to total variance, neither controlled for respondent age (which was reported as not varying across interviewers) nor examined interviewer effects for possible age differences. To our knowledge, there have been only two studies of general interviewer effects which directly examined the relationship between respondent age and interviewer variance. One such study is that reported by Freeman and Butler (1976), which, although it examined age of respondents only in combination with age of interviewers, found that "old" interviewers (32 and older) interviewing "old" respondents did indeed produce consistently more interviewer variance than did any other age combination. The other such study was done by Groves and Magilavy (1980), who found that interviewer effects were stronger for respondents age 60 or older than for young (age 18-24) respondents, for each of six types of questions that they examined.

## Data

The major problem in the study of interviewer effects consists of separating such effects from true differences between groups of persons interviewed by different interviewers. In personal interview studies, it is generally not practical to assign interviewers to sampled households on a random basis; for national surveys in particular, the travel costs and delays implied by such a procedure would be prohibitive. It is, however, possible to have "interpenetrating" designs, for example, one in which two or more interviewers are assigned randomly to sampled households within each geographic area, which permit estimation of interviewer effects (Mahalanobis, 1946; U.S. Census Bureau, 1979). Estimation of interviewer effects is considerably more feasible with telephone interviews, where typically all interviews are conducted from a central location. Of course, there is little basis for regarding estimates of interviewer effects obtained from a telephone survey to be unbiased estimates for interviewer effects in personal interview surveys; indeed, there is evidence that interviewer effects tend to be smaller in the telephone methodology (Groves and Magilavy, 1980).

The data used in the analyses to be reported below were obtained in a study by Cannell and Groves (1979a), using a computer-assisted telephone interviewing system. One feature of this system was that assignments of sampled telephone numbers to interviewers were generally made by the computer, initially on a random basis, then, after a number had been dialed without a response, in accordance with an algorithm that attempted to maximize the probability of a contact by taking account of the time of day and week. Since interviewers generally worked the same hours of day throughout the study, randomization could only be within each shift rather than across all interviewers. Moreover, some households that initially refused to participate in the study were then assigned to interviewers who had shown greater success in converting such refusals, but data from such interviews were deleted from our analyses.

The survey used a two-stage stratified sample of telephone numbers generated using the random digit dialing technique developed by Waksberg (1978). In an effort to reduce nonresponse bias, a double sampling scheme was used for nonrespondents after the fifth week of the data collection. An

interview was conducted with one objectively selected adult in each sample household. A total of 1054 interviews were completed, which represent a response rate of 67 percent if all telephone numbers which were never answered are included in the denominator (an unknown proportion of such numbers are non-working, or not located in households). The thirty or so persons who conducted these interviews were selected from applicants for the job in Ann Arbor; none had previous interviewing experience with the Survey Research Center, but all completed a three week training period. Most were female (87 percent), below thirty years of age (77 percent), and had completed at least two years of college (73 percent); about half were part time students (53 percent).

## Methods

Most studies of interviewer effects have utilized analysis of variance techniques, the independent variable being the interviewer identification, the dependent variables the responses to questions for which interviewer effects are being measured. In our analyses, we employed a regression analysis approach; this provides somewhat different statistics than does the analysis of variance approach to the study of interviewer effects, so we will provide our rationale for using this technique and examine the comparability of our measures with those from an analysis of variance.

The reason we chose to use regression analysis rather than analysis of variance is that regression analysis is the more flexible technique, allowing greater complexity to be incorporated more easily into the design matrix. Multiple and continuous variables (i.e., variables with many categories) with linear relationships to the dependent variable are easily included as predictors, whereas they are not easily handled (except as covariates) in analyses of variance. In the analysis of interviewer effects, for example, it might be interesting to use background characteristics of the respondents, such as education, age, and income, along with the interviewer dummy variables as predictors in order to examine interviewer characteristics as part of a broader explanatory model. In the analyses to be presented below, we included age and interviewer variables to examine their separate and interactive effects. The predictor variables in the regression analysis included a dummy variable for age (whether or not the respondent was 60 years old or older), a set of dummy variables representing all but one of the thirty interviewers, and another set of 29 dummy variables for the interaction of age and interviewers. This regression analysis is formally equivalent to an analysis of variance with two independent variables: a two category age variable, and a thirty category interviewer variable. One reason for preferring the use of regression analysis even in this case, despite its formal equivalence to analysis of variance, is that the standard output from regression analysis includes a set of coefficients each of which represents the effect of one interviewer. These coefficients can be compared for different dependent variables to determine whether certain interviewers consistently have large effects on the way respondents reply to their questions.

To measure the magnitude of interviewer effects, we have used the adjusted $R^2$ statistic, which is defined as:

$$R^2_{adj} = 1-[(N-1)/(N-c)](1-R^2),$$

where N is the total sample size and c is the number of interviewers.[1]

## Findings

### Effect of interviewer and age of respondent

The first question we address is whether interviewers tend to have the same effect on older respondents as they do on younger respondents. We used a set of items for which Groves and Magilavy (1980) found significant overall interviewer effects. Us ng an analysis of variance approach, these investigators tested a total of 77 items for interviewer effects, and found statistically significant effects ($p<.05$) for a total of 30 of these items, 14 of which are answers given by respondents, the other 16 observations made by the interviewers. We tested whether the age of the respondent's interacted with the specific interviewers with respect to each of those 30 items.

As noted before, in our regression analyses, dummy variables were included for all but one of the 30 interviewers involved in the data collection, an analysis that is formally equivalent to an analysis of variance in which the 30 interviewers constitute categories on the single independent variable. F-tests for the overall significance of interviewer effects are equivalent for the two types of analysis, but our measure of the magnitude of interviewer effects differs somewhat from that used in the analyses reported by Groves and Magilavy (1980:Table 1). The results are reported in Table 8-1. The first column lists the 30 variables, the second column reproduces the roh statistic reported by Groves and Magilavy, while the third column gives the adjusted $R^2$ values from our regression analysis when only the interviewer dummy variables are used as predictors. Comparing the two sets of values in columns 2 and 3, it is clear that they are very similar to one another.

The last three columns of Table 1 provide findings concerning the existence of interaction effects between interviewers and age of the respondents. Specifically, the fourth column lists the proportion of variance explained by interviewers and the dichotomous age variable, the "main effects" model. The next column lists the proportion of variance explained by a model that adds interaction terms for these two variables. (Note that these $R^2$ values are adjusted for degrees of freedom, so that the values in the fifth column can be, and sometimes are, smaller than those in the fourth column.) For eight of the thirty variables, there are statistically significant interaction effects, as shown by the F-statistic listed in the last column. These tests of significance should be regarded with some caution, since there is evidence that the variability of answers obtained by some interviewers is considerably larger than those obtained by other interviewers, a violation of the assumption of homoscedasticity which underlies the F-test. Nevertheless, it

---

[1]A statistic that has more frequently been used to measure interviewer effects is the "ratio of homogeneity," roh. As shown in the Appendix, if the number of interviewers c, is fairly large, both roh and $R^2_{adj}$ must both lie within narrow bounds, and thus must have similar values.

does appear that for a substantial proportion of variables, among those for which there are significant interviewer effects, there are also interactions between interviewers and the age of the respondents. In terms of explanatory power, comparison of the fourth and fifth columns of Table 1 indicates that the interaction effect explains between 1.5 and 4.5 percent of the variance beyond that explained by the main effects of age and interviewers.

As noted before, the last 16 variables listed in Table 1 are observations made by the interviewers after the completion of the interviews rather than responses given directly by the respondents. Of the 77 items that Groves and Magilavy examined for interviewer effects 19 were interviewer observations and 58 were answers by respondents to questions. Clearly, the interviewer observations were much more likely to be subject to interviewer effects than were the answers given by respondents: 16 out of the 19 observations were found to have statistically significant interviewer effects, compared to 14 out of the 58 responses. However, as shown in the analyses reported here, these interviewer observations were no more likely to show significant age by interviewer interaction effects than were the responses given directly by the respondents: 5 of the 14 responses have statistically significant interaction effects, as compared to 3 of the 16 interviewer observations, as shown in Table 1. We conclude that interviewers are more likely to differ systematically in the way they record their own observations than in the way they affect the answers given to them by respondents; but that age-by-interviewer interactions seem no more likely in the case of these interviewer observations than in the case of answers given by respondents.

Groves and Magilavy also reported that responses to open-ended questions tended to be subject to considerably larger interviewer effects than responses to closed-ended questions. If anything, the opposite pattern may hold with respect to age-by-interviewer interaction effects, however. The first six items listed in Table 1 are closed-ended, and for three of them there is a statistically significant interaction effect, whereas only two out of the eight open-ended questions showed significant interaction effects. In short, there is no evidence that age-by-interviewer interactions are more likely for open-ended than for closed-ended questions.

The analysis just described tells us that for a substantial proportion of variables which have significant interviewer effects, there are also significant interactions of interviewers and the age of the respondents, but it tells us nothing about the nature of those interactions. In particular, it does not address the primary question which motivated our investigation: whether interviewers tend to have more effect on answers given by older respondents than on those given by younger persons. To answer this question directly, we performed parallel regression analysis on two groups of respondents: those under the age of 60, and those age 60 or older. The magnitude of interviewer effects were estimated for each of the 30 variables within each age group. These magnitude estimates were described in the preceding section, and are adjusted to take account of the degrees of freedom and the size of the sample; thus the adjustments are considerably larger for the older group than for the younger group, since there were only 158 respondents age 60 or older compared to 785 under the age of 60.

The interviewer effects on each of the 30 variables and for each age group are shown in Table 2. The average interviewer effect (i.e., adjusted

$R^2$) for those age 60 or older is 7.5 percent, which is slightly higher than the average effect for younger respondents (7.0 percent). Since the distribution of values of the adjusted $R^2$ values are considerably skewed to the right, a more meaningful measure of the central values may be the medians: 6.2 percent for older respondents, which is twice as large as the median of 3.1 percent for younger respondents. As noted earlier, most of the variables for which there are large interviewer effects are interviewer observations rather than questions answered by the respondents. If we look only at the questions that were asked of respondents (i.e., the first 14 items in Table 2), the median proportion of explained variance is 5.0 percent for the older respondents, which is more than twice as large as the median of 2.0 percent observed for younger respondents. It appears that the proportion of variance in answers given by older respondents is greater than in the case of younger respondents. On the other hand, for the 16 interviewer observations listed at the end of Table 2 there is no evidence that there are stronger interviewer effects for older than for younger respondents: the median percent of explained variance is 8.0 percent for older respondents, which is actually less than the median of 10.8 percent for younger respondents, and for only 7 of the 16 items is the explained variance larger for the older than for the younger respondents. This observed difference for the two types of questions seems intuitively reasonable, since interviewer ratings are more exclusively under the control of the interviewer, while answers by respondents are a result of the interaction between the respondent and the interviewer.

## Checks on the Generality of the Findings

Importance of "Outlying" Interviewers. It is important to know whether the interviewer effects that we have observed reflect a small proportion of interviewers who rather consistently obtain answers from respondents that differ markedly from those obtained by other interviewers, or whether they represent a "normal" range of effects due to all of the interviewers. If a minority of interviewers were responsible, it might be possible to identify them and either retrain them or remove them from the interviewing staff. If some interviewers have relatively strong effects on certain types of questions (but not on others), it might be possible to identify interviewer characteristics that are associated with positive or negative biases on specific questions and to correct for those biases in the analysis.

Our first approach to this question was to examine the distributions of the regression coefficients (i.e., the deviations of the mean response obtained by each interviewer from the overall mean response to a question) estimated for each of 13 dependent variables. These variables represented answers by respondents rather than interviewer observations. The coefficients for each of these variables were examined for each of the 28 interviewers who interviewed both younger and older respondents, for a total of 364 coefficients for each age group.

Almost all of the coefficients with large absolute values (specifically, these that deviated from the overall mean for a dependent variable by at least two standard deviations) were associated with a few interviewers who obtained only a small number of interviews. The estimated effects of such interviewers are obviously less precise than for more productive interviewers; this shows up in the form of larger standard errors for the regression coeffi-

cients corresponding to those interviewers in the regression analysis. To determine whether this set of interviewers was responsible for the significant interviewer effects and/or for the significant age-by-interviewer interaction effects, we repeated these regression analyses omitting the dummy variables associated with all interviewers who failed to obtain at least five interviews with persons under the age of 60 and at least three interviews with older persons--a total of eight out of the thirty interviewers. For this analysis, we considered only the set of questions asked of respondents for which Groves and Magilavy found statistically significant interviewer effects--the first fourteen variables listed in Table 1.

These alterations did affect the findings, but not in what we consider a major way.[2] Our conclusion is that while some of the variance attributed to interviewers reflects the influence of a small number of interviewers who interviewed only a few respondents, a substantial amount of interviewer variance is due to a fairly normal distribution of effects associated with individual interviewers.

Consistency of Interviewer Effects across Questions. We next asked whether the effects attributed to individual interviewers are consistent across questions; that is, whether certain interviewers tend to bias answers given to them in a particular direction. This analysis was done for all interviewers and for all thirty of the items listed in Table 8-1, but for simplicity we only report our findings from an analysis restricted to items asked of respondents (omitting interviewer observations) and to interviewers who took at least a few interviews with both younger and older respondents (omitting the eight interviewers described in the preceding section).

The average correlation of the interviewer effects on responses to different questions among younger respondents was 0.242; for older respondents, the average was 0.069. For some of the items the interviewer effects are rather substantially correlated. Cluster analysis was used to identify a subset of nine items with respect to which the average correlation of the interviewer coefficients for younger respondents is 0.346.[3] This suggests that there may indeed be a consistent tendency for interviewers to obtain answers that are consistently biased in one direction or another. The correlations of

---

[2]The overall proportion of variance explained by interviewers (the average of the adjusted R-square values) dropped by about a third, from 2.4 to 1.6 percent, and interviewer effects were no longer significant at the .05 level for three of the fourteen variables. Statistically significant ($p<.05$) age-by-interviewer interaction effects were detected for four of the fourteen variables; two that are noted as being significant in Table 1 (V1013 and V1016) are no longer significant at this level when the eight interviewers are omitted, while a statistically significant ($p<.05$) interaction is now found for one variable (V1005) for which none was detected when all interviewers were considered. Moreover, inspection of the estimated regression coefficients revealed only 4.3 percent that differed by two or more standard deviations from the unweighted mean across all of the remaining interviewers.

[3]These nine items are, using the variable numbers listed in Table 1, V815, V924, V926, V928, V1005, V1025, V875, V1013, and V1024. We have not, however, tried to attach any meaning to this cluster of items.

the coefficients for the older respondents are considerably lower, however, perhaps because the number of older respondents in this study was rather small thereby yielding unreliable estimates of the effects of individual interviewers. It is also interesting that there is essentially no relationship between the relative effect of an interviewer on younger respondents and his or her effect on older respondents; the average correlation of the estimated effects across the items is only 0.089. For some of the items[4] the correlations are in the range from .40 to .50, but for other items the correlations are actually negative.[5]

## Summary and Conclusions

Using data collected through telephone interviews with a national sample of adults, we have searched for evidence as to whether or not interviewers have stronger effects on the responses given to a wide range of questions by older people than on the responses of younger people. Specifically, we examined responses to a set of thirty items for which significant interviewer effects had earlier been reported, out of a total of 77 items that were screened for such effects.

Significant age-by-interviewer interactions were detected for eight of the 30 items. The number of items is too small to permit conclusions to be drawn with any certainty about types of items that are more or less likely to be subject to such interactive effects. The set of eight items for which significant interactions were detected in this data set included both interviewer observations and the direct responses of interviewees, and among the latter responses to both open- and closed-ended questions.

We next examined older and younger respondents (i.e., those age 60 and older vs. all others) separately to determine the nature of the age-by-interviewer interactions. No pattern could be found for the interviewer observations, but for the direct responses to questions the median interviewer effect was more than twice as strong for older respondents as for younger respondents.

Some of the interviewer effects that we observed can be attributed to a small number of "deviant" interviewers who obtained answers quite different from those obtained by most other interviewers; these interviewers usually obtained only a small number of completed interviews, so that our estimates of their effects on the respondents are less reliable than for those who were more productive. However, removing these interviewers from the analysis did not change the basic pattern of the findings. We conclude that the observed age-by-interviewer effects probably reflect a range of interviewer characteristics and behaviors rather than a minority of poorly trained or deviant interviewers. With the present data, however, we can say nothing about what might be the nature of such characteristics.

---

[4]Variables numbered V815 (r = .41), V950 (r = .41), V1025 (r = .40), and V1024 (r = .49).

[5]Variables numbered V926 (r = -.07), V928 (r = -.08), V946 (r = -.03), and V1005 (r = -.09).

We also looked for evidence of consistency of the patterns of interviewer effects across different questions. For a subset of the questions answered by younger respondents we did find some evidence for at least a small degree of consistency, but no such consistency was found in the effects on older respondents. Moreover, the estimates of the effects of an interviewer on the answers given by younger respondents to a particular question were essentially unrelated to the estimates of the effects of that interviewer on the answers to the same question given by older respondents.

Based on the evidence that we have presented, it appears that older respondents may be somewhat more susceptible to interviewer effects than younger respondents. This implies that careful training of the interviewing staff is even more important when the target population is elderly than is true for general population surveys. Techniques for introducing standardized methods of asking questions and probing for clarification or amplification of answers should be implemented whenever possible, and in perticuler when th e is an emphasis on older persons. Procedures of instruction, commitment, and feedback developed by Cannell et al. (1981) and described elsewhere in this report have had the effect of reducing interviewer variance and might therefore be advisable for interviews with older adults.

Another implication of these findings is that the effective size of samples of older respondents may be considerably smaller than the apparent sample size. Interviewer effects, in a manner analogous to the more familiar design effects associated with complex sampling frames, have the consequence of increasing the magnitude of the errors associated with any sample statistic and thereby of reducing the effective sample size. This implies, for example, that the nominal probability levels cited for tests of hypotheses may be too lenient, perhaps leading to acceptance of hypotheses that are incorrect and thus, over the long run, to more frequent failures of attempts to replicate findings from previous studies.

These conclusions should be qualified, since our analyses of interviewer effects were based on data collected through telephone interviewing and we cannot be certain that similar patterns would be found in data collected by face-to-face interviews. An important difference between the typical telephone survey and the typical face-to-face survey is the number of interviews obtained by each interviewer; telephone interviewers, not restricted by geography, often collect several times as many respondents as do face-to-face interviewers who are confined to a small geographic area. Since the magnitude of the interviewer effects is proportional to the average number of respondents per interviewer, we would expect that if the average magnitude of the effect of an individual interviewer were about the same for the two interviewing modes, the overall interviewer effect would be considerably larger for telephone interview surveys than for face-to-face surveys.

On the other hand, there is little basis for thinking that individual interviewers would have any less influence on the respondents when they are face to face, and therefore perhaps more likely to communicate by non-verbal cues, than when they are in contact only by a telephone line. Indeed, Collins (personal communication, 1982) analyzed data from a face-to-face interview survey conducted in England, and found a similar pattern of interviewer-by-age of respondent interactions to those that we have discussed here. Moreover, the data used in our analysis were collected by a staff of interviewers in a

central location, where they were monitored throughout the data collection period, whereas the field staff in a face-to-face interview survey may be scattered over the entire country and feedback on their performance from the research staff is often necessarily very slow if not entirely lacking. For this reason we suspect that there is probably greater uniformity in the techniques and styles used by telephone interviewers than is typically true of a staff of face-to-face interviewers, and this leads us to expect that interviewer effects are even stronger for face-to-face interview surveys than those we found for this telephone interview study. Whether interviewer-by-age of respondent interactions may also be more frequent in face-to-face interview studies than we observed in this study would be purely speculative; but we see little basis for expecting such interactions would be any less frequent in face-to-face interviews than in telephone interviews.

Pending the collection of more data that would permit the direct evaluation of the frequency and magnitude of interviewer effects on older respondents, our conclusion is that caution should be applied in the interpretation of data obtained by face-to-face interviews with the elderly. Interviewer effects may be substantial, and could distort the apparent probabilities of sample statistics; and could introduce biases into such statistics, since there is no guarantee that the average effect across a set of interviewers is zero. This caution should be particularly heeded when the number of respondents interviewed by each interviewer is substantial.

## APPENDIX

The "ratio of homogeneity," or roh, which is defined (cf. Kish, 1965:172) as follows:

$$roh = (s_a^2 - s_b^2/b)/s^2,$$

where $s_a^2$ is the "between interviewer" component of the variances:

$$s_a^2 = [1/(a-1)] \sum_{i=1}^{a} (\bar{y}_i - \bar{y})^2$$

$$= SSR/[b(a-1)] = (SSTO - SSE)/[b(a-1)];$$

$s_b^2$ is the "within interviewer" component of the variance:

$$s_b^2 = 1/[a(b-1)] \sum_{i=1}^{a} \sum_{j=1}^{b} (y_{ij} - \bar{y}_i)^2$$

$$= SSE/[a(b-1)]; \quad \text{and}$$

$$s^2 = s_a^2 + [(b-1)/b]s_b^2$$

$$= SSTO/[b(a-1)] - SSE/[ab(a-1)];$$

with a being the number of interviewers and b being the number of interviews taken by each interviewer.[4] In our regression analyses, the measure of interviewer effects is the adjusted $R^2$ value, defined as follows:

$$R_{adj}^2 = 1 - [(N-1)/(N-c)](1-R^2).$$

This can be rewritten as follows (since N = ab and c = a):

$R_{adj}^2 = 1 - [(ab-1)/(ab-a)]SSE/SSTO$. A little algebra shows us that both roh and the adjusted $R^2$ value must lie within the following bounds:

$$1 - [b/(b-1)]SSE/SSTO \leq \begin{bmatrix} roh \\ R_{adj}^2 \end{bmatrix} \leq 1 - [(a-1)/a] * [b/(b-1)]SSE/SSTO,$$

---

[4] In practice, the number of interviews taken by each interviewer is not a constant; in the present survey, one of the 30 interviewers took only 4 interviews, while another took more than 80 interviews. The expression for roh should be modified in terms of the average number of interviews per interviewer; see Groves and Magilavy (1980). The sampling distribution of roh is also effected by variability in numbers of interviews per interviewer.

## REFERENCES


Bailar, B.A. Some sources of error and their effect on census statistics. Demography, 1976, 13, 273-286.

Bailar, B.A., Bailey, L., and Stevens, J. Measures of interviewer bias and variance. Journal of Marketing Research, 1977, 14, 337-343.

Bailey, L., Moore, T.F., and Bailar, B.A. An interviewer variance study of the eight impact cities of the National Crime Survey cities sample. Journal of the American Statistical Association, 1978, 73, 16-23.

Benney, M., Gray, A.P., and Pear, R.H. How people vote. New York: Grove Press, 1956.

Boyd, H.W., Jr., and Westfall, R. Interviewers as a source of error in surveys. Journal of Marketing, 1955, 19, 311-324.

Boyd, H.W., Jr., and Westfall, R. Interviewer bias re-visited. Journal of Marketing Research, 1965, 2, 58-63.

Boyd, H.W., and Westfall, R. Interviewer bias once more re-visited. Journal of Marketing Research, 1970, 7, 249-253.

Cahalan, D. Correlates of respondent accuracy in the Denver validity survey. Public Opinion Quarterly, 1968, 32, 607-621.

Campbell, A., Converse, P.E., and Rodgers,W.L. The quality of American life. New York: Russell Sage, 1976.

Cannell, C.F., and Groves, R.M. Development and experimentation with telephone data collection techniques [machine-readable data file]. Ann Arbor, MI: Survey Research Center, Institute for Social Research [producer], 1979(a).

Dohrenwend, B.S., Colombotos, J., and Dohrenwend, B.P. Social distance and interviewer effects. Public Opinion Quarterly, 1968, 32, 410-422.

Ehrlich, J.S., and Riesman, D. Age and authority in the interview. Public Opinion Quarterly, 1961, 25, 39-56.

Freeman, J., and Butler, E.W. Some sources of interviewer variance in surveys. Public Opinion Quarterly, 1976,

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

Groves, R.M., and Magilavy, L. Estimates of interviewer variance in telephone surveys. Proceedings of the 1980 American Statistical Association Meeting. Washington, DC: American Statistical Association, 622-627.

Hanson, R.H., and Marks, E.S. Influence of the interviewer on the accuracy of survey results. Journal of the American Statistical Association, 1958, 53, 635-655.

Herzog, A.R. Attitude change in older age: An experimental study. Journal of Gerontology, 1979, 34, 697-703.

Herzog, A.R., and Rodgers, W.L. Age and Satisfaction: Data from several large surveys. Research on Aging, 1981, 3, 142-165. (a)

Hyman, H.H., Cobb, W.J., Feldman, J.J., Hart, C.W., and Stember, C.H. Interviewing in Social Research. Chicago: University of Chicago Press, 1954, Chapter 6.

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

Klein, R.L. Age, sex, and task difficulty as predictors of social conformity. The Journal of Gerontology, 1972, 27, 229-236.

Klein, R.L., and Birren, J.E. Age differences in social conformity on a task of auditory signal detection. Proceedings of the 80th Annual Convention, American Psychological Association, 1972, 661-662.

Mahalanobis, P.C. Recent experiments in statistical sampling in the Indian Statistical Institute. Journal of the Royal Statistical Society, 1946, 109, 326-378.

McKenzie, J.R. An investigation into interviewer effects in market research. Journal of Marketing Research, 1977, 14, 330-336.

Stock, J.S., and Hochstim, J.R. A method of measuring interviewer variability. Public Opinion Quarterly, 1951, 15, 322-334.

Sudman, S., and Bradburn, N.M. Response effects in surveys: A review and synthesis. Chicago: Aldine, 1974.

Sudman, S., Bradburn, N.M., Blair, E., and Stocking, C. Modest expectations: The effects of interviewers' prior expectations on responses. Sociological Methods and Research, 1977, 6, 177-182.

U.S. Bureau of the Census. Enumerator variance in the 1970 census, PHC(E)-13. June, 1979a.

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

Table 1: Interviewer, Age, and Interaction Effects on

Selected Variables from the Groves-Cannell Telephone Survey

| Var. # | Variable Name | roh | $R^2$ adj. | | | F test (interaction) |
|---|---|---|---|---|---|---|
| | | | INT | INT,AGE | INTxAGE | |
| V815 | I watch television because there is nothing else to do: hardly ever | .018 | .0170 | .0227 | .0178 | 0.836 |
| | Proportion ever having had: | | | | | |
| V924 | Arthritis or rheumatism | .023 | .0218 | .1202 | .1487 | 2.100# |
| V926 | Heart disease or any heart trouble | .021 | 0201 | .0438 | .0850 | 2.475# |
| V928 | Hypertension or high blood pressure | .018 | .0167 | .0689 | .0911 | 1.803+ |
| V932 | Proportion having made contributions to the American Heart Association | .022 | .0210 | .0523 | .0544 | 1.069 |
| V946 | In the previous month drank hard liquor: proportion not at all | .016 | .0145 | .0224 | .0119 | 0.646 |
| V947 | Time elapsed since last doctor's visit, less than 6 months | .027 | .0257 | .0294 | .0330 | 1.115 |
| V950 | Proportion able to report date of most recent doctor's visit, precise within a week | .031 | .0306 | .0299 | .0248 | 0.836 |
| V1005 | Number of health conditions reported in open questions | .016 | .0152 | .0467 | .0586 | 1.418 |
| V1025 | Number of ameliorative health behaviors reported | .028 | .0271 | .0348 | .0327 | 0.930 |
| V875 | Number of ways mentioned that TV is good or bad for children | .016 | .0146 | .0163 | .0128 | 0.887 |
| V1013 | Length of response for mentions of "bad reactions" to medicine | .024 | .0230 | .0244 | .0425 | 1.621* |
| V1016 | Length of response for mentions of health symptoms | .031 | .0298 | .0318 | .0520 | 1.698* |
| V1024 | Length of response for mentions of ameliorative health behaviors | .061 | .0581 | .0651 | .0595 | 0.804 |
| | Interviewer Observations: | | | | | |
| V902 | Proportion of respondents "never" giving an answer so long that the interview was delayed by typing | .241 | .2329 | .2324 | .2385 | 1.253 |
| V909 | Proportion of interviewers having a problem with the terminal control keys | .047 | .0446 | .0449 | .0785 | 2.168# |

| Var # | Variable Name | roh | $R^2$ adj. | | | F test (interaction) |
|---|---|---|---|---|---|---|
| | | | INT | INT,AGE | INTxAGE | |
| V912 | Proportion of interviewers reporting that there were "Sometimes" slow system response times | .104 | .0998 | .0988 | .0991 | 1.011 |
| | Did respondent give an indication of checking any records to look up last doctor's visit? | | | | | |
| V803 | Proportion uncertain | .237 | .2295 | .2288 | .2145 | 0.399 |
| V804 | Proportion no | .165 | .1593 | .1584 | .1605 | 1.083 |
| V878 | Proportion of respondents "not at all" reluctant to agree to interview | .055 | .0537 | .0681 | .0735 | 1.123 |
| V881 | Proportion of respondents "not at all" suspicious about survey | .087 | .0834 | .0965 | .1125 | 1.579* |
| V882 | Proportion of respondents whose suspicion was reduced as the interview progressed | .180 | .1710 | .1871 | .1658 | 0.785 |
| V885 | Proportion of respondents who seemed to rush their answers | .025 | .0237 | .0234 | .0312 | 1.258 |
| V886 | Proportion of respondents who did not ask how much longer the interview would last | .023 | .0223 | .0548 | .0688 | 1.482 |
| V890 | Proportion of respondents who seemed "somewhat interested in the interview | .036 | .0342 | .0334 | .0424 | 1.299 |
| V892 | Proportion of respondents who "never" asked for clarification during the interview | .186 | .1794 | .1804 | .1873 | 1.273 |
| V896 | Proportion of respondents who needed "a few question" repeated | .092 | .0869 | .0885 | .1333 | 2.657# |
| V898 | Proportion of respondents suspected of not giving honest answers | .021 | .0198 | .0205 | .0088 | 0.621 |
| V903 | Proportion of respondents making "very much effort" to give complete and accurate answers | .152 | .1462 | .1454 | .1451 | 0.988 |
| V906 | Proportion of respondents "very likely" to grant a re-interview | .096 | .0930 | .0990 | .1073 | 1.295 |

* $p < .05$ + $p < .01$ # $p < .001$

Table 2: Interviewer Effects on Younger (Under age 60) and Older Respondents

| Var. # | Variable Name | $R^2$ adj. (Age < 60) | $R^2$ adj. (Age ≥ 60) |
|---|---|---|---|
| | Respondent Answers: | | |
| V815 | I watch television because there is nothing else to do: hardly ever | .0049 | .0901 |
| | Proportion ever having had: | | |
| V924 | Arthritis or rheumatism | .0108 | .0547 |
| V926 | Heart disease or any heart trouble | .0186 | .0679 |
| V928 | Hypertension or high blood Pressure | .0212 | .0304 |
| V932 | Proportion having made contributions to the American Heart Association | .0169 | .0321 |
| V946 | In the previous month drank hard liquor: proportion not at all | .0123 | (-.0428) |
| V917 | Time elapsed since last doctor's visit, less than 6 months | .0219 | .0543 |
| V950 | Proportion able to report date of most recent doctor's visit, precise within a week | .0220 | .0463 |
| V1005 | Number of health conditions reported in open questions | .0163 | (-.0072) |
| V1025 | Number of ameliorative health behaviors reported | .0184 | (-.0013) |
| V875 | Number of ways mentioned that TV is good or bad for children | .0247 | (-.0892) |
| V1013 | Length of response for mentions of "bad reactions" to medicine | .0228 | .1033 |
| V1016 | Length of response for mentions of health symptoms | .0431 | .0492 |
| V1024 | Length of response for mentions of ameliorative health behaviors | .0525 | .0387 |
| | Interviewer Observations: | | |
| V902 | Proportion of respondents "never" giving an answer so long that the interview was delayed by typing | .2475 | .1793 |
| V909 | Proportion of interviewers having a problem with the terminal control keys | .0538 | .4781 |

| Var. # | Variable Name | $R^2$ adj. (Age < 60) | $R^2$ adj. (Age ≥ 60) |
|---|---|---|---|
| V912 | Proportion of interviewers reporting that there were "sometimes" slow system response times | .1122 | .0228 |
| | Did respondent give an indication of checking any records to look up last doctor's visit? | | |
| V803 | Proportion uncertain | .2002 | .2916 |
| V804 | Proportion no | .1445 | .2423 |
| V878 | Proportion of respondents "not at all" reluctant to agree to interview | .0342 | .0993 |
| V881 | Proportion of respondents "not at all" suspicious about survey | .0983 | .0827 |
| V882 | Proportion of respondents whose suspicion was reduced as the interview progressed | .2392 | (-.1584) |
| V885 | Proportion of respondents who seemed to rush their answers | .0284 | .0222 |
| V886 | Proportion of respondents who did not ask how much longer the interview would last | .0269 | (-.0297) |
| V890 | Proportion of respondents who seemed "somewhat" interested in the interview | .0415 | .0381 |
| V892 | Proportion of respondents who "never" asked for clarification during the interview | .1602 | .3578 |
| V896 | Proportion of respondents who needed "a few question" repeated | .1318 | .1007 |
| V898 | Proportion of respondents suspected of not giving honest answers | .0163 | (-.0809) |
| V903 | Proportion of respondents making "very much effort" to give complete and accurate answers | .1483 | .0677 |
| V906 | Proportion of respondents "very likely" to grant a re-interview | .1033 | .0632 |