

## THE EFFECTS OF INTERVIEWER CHARACTERISTICS ON ITEM RESPONSE

John D. Hutcheson, Jr., James E. Prather and Rob E. Snow, Georgia State University

The purpose of this paper is to identify interactions between interviewer and respondent characteristics that bias or inhibit responses to different types of schedule items. While the problem of refusals to be interviewed, and the personnel and schedule format factors which influence refusals have been widely explored, the interactions of interviewer and respondent characteristics, and their impact on responses to particular types of items have also received considerable attention. Aspects of this topic are a recurring theme in the literature on conducting survey research (Bauman, 1971; Cable, 1971; Colombotos, 1968; Glasser and Metzger, 1970; Hawkins and Cable, 1970; Hyman et al., 1954; Lansing, 1971; McClelland, 1974; Price and Searles, 1961; Robins, 1974; Singer and Kohnke-Aguirre, 1979; Sudman and Bradburn, 1974; Sudman, Bradburn and Blair, 1977; Veltman, 1972; Waller, 1968; Weiss, 1975; and Williams, 1964, 1968). To the extent that interviewer characteristics intrude on the measurement process, valuable data are being lost or distorted (Barr, 1957; Boyd and Westfall, 1955, 1964, 1967; Duncan, 1979; Feldman et al., 1951; Hill, 1973; Hanson and Marks, 1958; Sudman and Bradburn, 1974). This paper seeks to identify potential impediments to adequate measurement by analyzing a large data set (n=7059, 400 items) compiled by forty interviewers to determine if systematic bias or problem responses (missing, refusals) are related to a range of interviewer and respondent background variables and measures of interviewer productivity and performance. The interviews occurred at randomly selected households in a large SMSA and were conducted by formally trained (3 days) and closely supervised professional interviewers. Specifically, this research focuses on the interaction of interviewer and respondent characteristics that maximize the probability of meaningful unbiased responses. In doing so it is hoped that the paper's findings might have wide application among practitioners in the survey research field. The data analyzed permit investigation of a number of combinations of interviewer and respondent traits upon response to several types of items.

In an ideal world, survey researchers would hope that the interactions between interviewer characteristics and respondent's answers to specific types of items would tend to cancel each other out and leave the responses unbiased, even if somewhat less efficient. On the other hand, if it can be demonstrated that a respondent is responding to items in a certain way because of the age, race, or sex of the interviewer, then the survey researcher's data typically contain an unspecified contamination due to interviewer characteristics rather than "noise" expected in an ideal world. One purpose of this paper is to identify some possible topic areas where an item might not yield valid answers due to the match of respondent and interviewer

characteristics (see Sudman and Bradburn, 1974). The survey data examined include responses to such sensitive questions as those concerning women's rights, personal safety, gun ownership, personal finance, receipt of public assistance, and religious preference. In addition, the structure of items in terms of the number of possible responses was investigated in relation to interactions of respondent-interviewer characteristics.

Data on the background, selection, training, productivity, and potential of interviewer were also available. This information is used to specify the above relationship and includes such performance measures as length of the interview, the number of calls necessary to complete the interview, the length of time the interviewer was employed, the date of the interview, total number of interviews conducted, refusals, and their evaluation by supervisors (see Cannel, 1970; Collins, 1970; Niamas, 1962, 1966; U.S. Bureau of the Census, 1972; and Williams, 1968). Data on selection and training included scores on a series of tests administered during the selection of interviewers, and a post-test administered after three days of training to gauge ability to handle item wording, follow instructions, record answers legibly, and speak audibly and clearly to respondents.\*

Before presenting the data analyses there are certain caveats that should be noted. This is an ad hoc secondary analysis of an existing data base. The experimental controls that are necessary to unambiguously address cause and effect are absent. On the other hand, this large data base does allow a rich analysis of complex interactions. Thus, it is necessary to note possible intervening factors and to be aware that interactions found may be due to spurious association attributable to a natural sorting process among interviewers and respondents on the basis of sex, race or age. Only an experiment could provide precise definition of these interactions.

\*The interviewers selected from the screening process were assigned to 1 of 8 teams composed of 4 to 6 interviewers and one team coordinator, most of whom had graduate degrees in social science disciplines. Training took place in small groups, and used simulated interviews with such aids as video tapes to illustrate interviewing skills such as probing and use of filter items. The interview schedule was pre-tested and an ongoing training program was maintained in order to deal with problems arising in the field. Completed interviews were verified by calling a 10% sample of respondents' households. So called "shade-tree" interviews were deleted from the data set and the interview site reassigned. Of course, not all falsifications could be caught with this method, but quality control was provided (see Evans, 1961).

## METHODS

Given the tentative nature of probing for the complex effects of the interactions among interviewers' and respondents' characteristics, the chosen method of analysis is one that starts at an exploratory level. It was thought that by presenting bivariate associations, and carefully looking for high-level relationships this paper could avoid the problems of multivariate analysis until a base had been established upon which complex models could be developed. Problems of multicollinearity are present since simple inspection of the list of interviewer characteristics clearly shows many common elements where one variable could easily be a proxy for another. Additionally, the functional form of multivariate equations (linear, log-linear, etc.) cannot be determined until such exploratory analysis has been completed.

Sensitive items in the survey were analyzed by presenting the interactions by respondent and interviewer for sex, race, and age. These interactions were defined by cross-classifying responses by the characteristics of the interviewer.\* For example, male respondents' answers were cross-classified by whether a male or female interviewed them and the same procedure was applied to female respondents. The respondent-interviewer age interactions were somewhat arbitrarily defined as whether:

1. the interviewer's age was within five years of the respondent's age;
2. the respondent was from 5 to 19 years older than the interviewer;
3. the respondent was 20 or more years older than the interviewer; and
4. the interviewer was six or more years older than the respondent.

The rationale for using the above classifications is based on questions of whether combinations of age, race, and sex roles were prompting or inhibiting responses. Thus, the responses on certain items for cross respondent-interviewer combinations on age, race, and sex were available to compare with the responses when the same age, race, and sex combinations for interviewer and respondent occurred.

To facilitate the interpretation of these cross classifications, a summary measure was used to gauge the effect of knowing the race, sex or age of the interviewer upon predicting the answer of a respondent of a given age, sex, or race category. The summary measure chosen is the uncertainty coefficient<sup>†</sup> which gives the

\*See Weiss (1975: 381-386) for an excellent review of the literature on the effects of the match between interviewer and respondent characteristics.

†The uncertainty coefficient describes how well the entire distribution of the dependent variable can be predicted on the basis of prior knowledge of the independent variables.

proportion by which "uncertainty" in the response (i.e., dependent variable) is reduced by simply knowing the age, sex, or race of the interviewer. Ideally, the uncertainty coefficient would be zero or trivially above zero. An uncertainty coefficient that was found to be "high" enough above zero would mean either:

1. the interviewer's race, sex, or age was influencing the respondent's answer; or
2. the interviewers of that particular sex, race, or age category were disproportionately and nonrandomly interviewing that category of respondent.

No guidelines exist of predetermining when how much is too much, but the uncertainty coefficient can alert one to make close inspection of possible interviewer induced-bias. It should be noted that the uncertainty coefficient's value is indeed contingent upon the marginals of a given table, so caution must be exercised in comparing uncertainty coefficients across tables.

In addition to illuminating the effects of the interviewer-respondent cross classifications by age, race, and sex on responses to particular types of items, also illustrated are the effects of interviewer productivity, background, actual performance and potential for performance. Of particular interest were the relationships of problem responses (missing, don't know) to these additional interviewer characteristics. Productivity was measured by length of interview, calls necessary to set it up, the interview date\*, number of days employed and refusals. Background information is incorporated by use of variables indicating interviewer age and years of formal education. Performance is gauged by the average of periodic evaluations by each interviewer's supervisor, and interviewer's potential for performance is measured by the screening test (with subscores) and with training posttest scores.

The response categories for the types of items were used for a comparison of mean scores on the above interviewer variables. To help in analyzing these means, a standard F-ratio for one-way analysis of variance is given. No significance test was done, nor are any inferential conclusions drawn because this is an exploratory study using a univariate measure on multivariate data. In addition, the means are not strictly independent in that a particular interviewer would, in all likelihood, have his/her score on a variable in more than one category. This lack of independence does affect the interpretation of the F-ratio since it violates the

\*Interview dates were used to specify whether an interview occurred early or late during the time period during which interviews were completed. The months of the interviews ranged from March (3) to September (9) and the days were converted to a decimal. For example, March 15 was coded as 3.5.

independence assumption; but for this paper's purpose of suggesting exploratory findings it should not present a major problem.

## FINDINGS

The findings from analyses of these data are presented by groups of variables representing categories of sensitive items about which responses may be influenced by the match between interviewer and respondent characteristics. In addition, quality control items that are based on interviewers' evaluations of respondent characteristics are presented. The categories used to organize the analysis include: women's issues, receipt of public assistance, public and personal safety, personal and financial questions. Data on financial and personal safety are presented here as example results. The analyses on the remaining types of items are available from the authors.

### Personal and Public Safety:

Table 1 contains an item relating to the respondent's perception of personal security. The responses to the item asking whether a respondent was afraid to walk about at night (Table 1) illustrate that male respondents admitted fear to female interviewers by answering "yes" 38.8% of the time. When the interviewer was male, the response was "yes" 31.3% of the time. Males were prone to report lower levels of fear to male interviewers than to female interviewers. Such differences could be attributed to male and female role expectations concerning machismo when male respondents were interacting with male interviewers. This is possibly a case where the respondent is giving a socially desirable answer (See Sudman and Bradburn, 1974:114).

### Personal and Financial Questions:

One of the most sensitive questions in many surveys concerns income. Table 2 shows that problem responses to this item are associated with respondent-interviewer race interactions. White respondents are less likely to tell a white interviewer (19.9%) than a black interviewer (32.0%) that they "don't know" their income, but white respondents also refuse to tell white interviewers (14.6%) more than black interviewers (9.6%). Unfortunately, this analysis suggests little that would increase response to the income question, given this response pattern among white respondents.

The religion question (Table 3) shows predicted patterns of religious preference, although black respondents were more prone to answer "none" to white interviewers (4.9%) than to black interviewers (2.7%). In this situation, black interviewers may have been less likely to accept a "none" response to the religion item from a black respondent. Black interviewers may have been willing to probe for religious socialization or affiliation with a

church as a child while white interviewers simply accepted an initial "none" response from black respondents. A lower rate of "none" as a response for the religion item was also found when white respondents were interviewed by black interviewers than when they were interviewed by white interviewers.

## SUMMARY AND CONCLUSIONS

The previous analyses of a large data set compiled by trained interviewers illustrates that there are interviewer-respondent age, race, and sex combinations that may have an impact on the responses elicited from a respondent. For example, white respondents and white interviewers may yield a disproportionate refusal rate when queried about income, yet when the interviewer is black a disproportionate rate of don't know responses, rather than refusals, are given. Or, in the case of reporting fear for personal safety, males were more prone to report lower level of fear to male interviewers than to female interviewers. These data are convergent with a number of studies that have previously documented that respondents tailor their answers in the direction of social desirability or stereotypic response when interviewers manifest particular age, race, sex, or social-class attributes. In addition, the data strongly support Sudman and Bradburn's conclusion that "...response interactions due to characteristics of respondents and interviewers are not general, but depend very specifically on the topic of the question" (1974: 123).

While these data present evidence that the interviewer-respondent interactions are not uniform across survey items, the current analysis was not an experiment which could clearly establish linkages among interviewer and respondent characteristics. Following the suggestions of Feldman (1951), and Sudman and Bradburn (1974) we also suggest that future research experimentally explicate the relationships among these variables. Given the lack of such experimental studies, the practitioner of survey research is enjoined to test for respondent-interviewer interaction effects when drawing conclusions concerning sensitive issues when using survey data.

## REFERENCES

- Barr, A., "Differences between experienced interviewers," Applied Statistics, Vol. 6, No. 3, (November 1957), pp. 180-189.
- Bauman, L. J., Rodgers, T. F., and Weiss, C. H., "Studies assessing the effects of matching between interviewer and respondent." In Abstracts of Papers on Respondent-Interviewer Interaction in the Research Interview, (April, 1971), pp. 12-53.
- Benney, M., Riesman, D., and Star, S. A., "Age and sex in the interview," American Journal of Sociology, Vol. 62, no. 6, (July 1956), pp. 140-147.

- Boyd, H. W., and Westfall, R., "Interviewer bias once more revisited," Journal of Marketing Research, Vol. 17, No. 2, (February 1970), pp. 249-253.
- \_\_\_\_\_, "Interviewer bias revisited," Journal of Marketing Research, Vol. 12, No. 1, (February 1965), pp. 58-63.
- \_\_\_\_\_, "Interviewers as a source of error in surveys," Journal of Marketing Research, Vol. 19, No. 4, (April 1955), pp. 311-324.
- Brodburn, N. M., Sudman, S., Blair, E., and Stocking, C., "Question threat and response bias," Public Opinion Quarterly, Vol. 42, No. 2, (Summer 1978), pp. 221-234.
- Cable, J. A., "Results of the Pilot Study Racial Match Experiment." In J. B. Lansing, et al., Working Papers on Survey Research in Poverty Areas, Ann Arbor: Institute for Social Research (University of Michigan), (January 1971), pp. 60-91.
- Cannell, C. F., "Interviewer performance over time: Some implications for field supervision and training." In C. F. Cannell, K. H. Marquis, and A. Laurent, Studies of Interviewing Methodology--A Summary of Research Conducted for the National Center for Health Statistics, (May 1970).
- Collins, A. W., "Verbal idiosyncrasies as a source of bias," Public Opinion Quarterly, Vol. 34, No. 3, (Fall 1970), pp. 416-422.
- Colombotos, J., Elinson, J., and Loewenstein, R., "Effect of interviewer's sex on interview responses," Public Health Reports, Vol. 83, No. 8, (August 1968), pp. 685-690.
- Duncan, R. P., "Survey quality and interviewer attributes," Southern Sociologist, Vol. 10, No. 3, (Summer 1979), pp. 16-21.
- Evans, F. B., "On interviewer cheating," Public Opinion Quarterly, Vol. 25, No. 1, (Spring 1961), pp. 126-127.
- Feldman, J. J., Hyman, H., and Hart, C. W., "A field study of interviewer effects on the quality of survey data," Public Opinion Quarterly, Vol. 15, No. 4, (Winter 1951), pp. 734-762.
- Glasser, G., and Metzger, G. A., "Measurement and control of interviewer variability." In Proceedings of the Business and Economic Section, American Statistical Association, (1970), pp. 314-317.
- Hanson, R. H., and Markes, E. S., "Influence of the interviewer on the accuracy of survey results," Journal of the American Statistical Association, Vol. 53, No. 283, (September 1958), pp. 635-655.
- Hawkins, L., and Cable, J. A., "The problem of response error in interviews." In J. B. Lansing, et al., Working Papers on Survey Research in Poverty Areas, Ann Arbor: Institute for Social Research (University of Michigan), (January 1971), pp. 60-97.
- Hill, R. J., and Hall, N. E., "A note on rapport and the quality of interviewer data," Southwestern Social Science Quarterly, Vol. 44, No. 3, (December 1973), pp. 247-255.
- 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), p. 415.
- Lansing, J. B., and Morgan, J. N., "Methods of data collection." In Economic Survey Methods, Ann Arbor: Institute for Social Research (University of Michigan), (May 1971), Chapter 4.
- McClelland, L., "Effects of interviewer-respondent race interactions on household interview measures of motivation and intelligence," Journal of Personality and Social Psychology, Vol. 29, No. 3, (1974), pp. 392-397.
- Namias, J., "A rapid method to detect differences in interviewer performance," Journal of Marketing, Vol. 26, No. 2, (April 1962), pp. 68-72.
- Price, D. O., and Searles, R., "Some effects of interviewer-respondent interaction on responses in a survey situation." In Proceedings of the Social Statistical Section, American Statistical Association, (1961), pp. 211-221.
- Robins, L. N., "Validity of interviews." In Executive Office of the President, The Vietnam Drug User Returns (Special Action Office for Drug Abuse Prevention Monograph, Series A., No. 2), Washington, D.C.: U.S. Government Printing Office, (May 1974).
- Singer, E., and Kohnke-Aguirre, L., "Interviewer expectation effects: A replication and extension," Public Opinion Quarterly, Vol. 43, No. 2, (Summer 1979), pp. 245-260.
- Stevens, J. A., and Bailer, B. A., "The relationship between various interviewer characteristics and the collection of income data." Presentation at the Meeting of the American Statistical Association, Boston, (1976).
- Suchman, H. and Converse, J. M., "The effects of Black and White interviewers on Black responses in 1968," Public Opinion Quarterly, Vol. 35, No. 1, (Spring 1971), pp. 44-46.
- Sudman, S., Reducing the Cost of Surveys, Chicago: Aldine Publishing Company, (1967), p. 246.
- Sudman, S., and Bradburn, N. M., "The effect of respondent and interviewer characteristics on response of attitudinal questions." In Response Effects in Surveys--A Review and Synthesis, Chicago: Aldine Publishing Company, (1974), Chapter 4.
- Sudman, S., Bradburn, N. M., Blair, E., and Stocking, C., "Modest expectations: The effects of interviewer's prior expectations on responses," Sociological Methods and Research, Vol. 6, No. 2, (November 1977), pp. 177-182.

U.S. Bureau of the Census, Investigation of Census Bureau Interviewer Characteristics, Performance, and Attitudes: A Summary (Working Paper No. 34), Washington, D. C.: U.S. Government Printing Office, (1972).

Veltman, C. J., "The resistance of respondents in inter-ethnic interviewing," Sociology and Social Research, Vol. 56, (July 1972), pp. 513-

Waller, L., and Fuchterland, E., "Interviewer-respondent interaction in Negro and White family life research," Human Organization, Vol. 27, No. 1, (Spring 1968), pp. 50-55.

Weiss, C., "Interviewing in evaluation research." In E. L. Struening and M. Guttentag (Eds.), Handbook of Evaluation Research (Vol. 1), Beverly Hills: Sage Publications, Inc., (1975), pp. 355-395.

Williams, A. J., "Interviewer-respondent interaction: A study of bias in the interview situation," Sociometry, Vol. 27, No. 3, (September 1964), pp. 337-347.

., "Interviewer role performance: A further note on bias in the information interview," Public Opinion Quarterly, Vol. 32, No. 2, (Summer 1968), pp. 287-294.

TABLE 1  
EFFECTS OF INTERVIEWER CHARACTERISTICS UPON RESPONSES  
TO ITEM:

"Is there any area right around here - that is, within half a mile - where you would be afraid to walk at night?"

	NA	Yes	No	DK	Coefficient	N
Total	1.4%	54.2%	43.3%	1.3%		
Interviewer 6+ Yrs. Older	1.0	50.3	47.7	1.0	.005	901
Interviewer Within 5 Yrs.	.5	52.3	46.4	.8		1171
Respondent 5-19 Yrs. Older	.8	50.3	47.7	1.2		1733
Respondent 20+ Yrs. Older	.9	58.7	38.8	1.6		3010
<u>Respondent</u>						
<u>Interviewer</u>						
Male	.8	38.8	59.9	.9	.004	1272
Male	1.1	31.3	66.7	.9		1590
Female	.9	68.6	28.5	2.0	.001	2137
Female	.9	67.8	30.1	1.1		1949
White	1.1	55.5	41.7	1.7	.004	1069
White	.5	50.6	48.3	.7		2286
Black	.7	57.4	40.4	1.5	.001	2983
Black	.9	54.7	43.0	1.4		428
		Means			F	
Interview Length (minutes)	44.37	58.55	58.07	59.57	13.03	
Calls Necessary	1.62	1.46	1.48	1.23	3.03	
Interview Date	6.02	6.45	6.65	6.45	20.40	
Interviewers: Age	32.51	30.91	31.09	32.60	1.13	
Schooling	15.44	15.32	15.33	15.26	.18	
Days Employed	73.40	70.03	69.78	64.62	1.35	
Total Interviews	183.50	181.90	177.60	177.30	1.73	
Refusals (number)	42.17	37.07	38.07	33.60	3.47	
Training Posttest	68.86	67.37	68.29	62.32	3.96	
Evaluation of Interviewer	38.25	37.34	37.78	37.13	2.20	
Screening Test: Total	37.86	37.71	37.70	37.90	.43	
Wording	8.87	8.67	8.63	8.91	2.93	
Following Instructions	9.29	9.27	9.30	9.39	.85	
Legibility	9.78	9.89	9.91	9.63	.59	
Audibility	9.86	9.97	9.97	9.96	.07	
N	91	3715	2946	89		

TABLE 2  
EFFECTS OF INTERVIEWER CHARACTERISTICS UPON RESPONSES TO  
ITEM: "About how much was your total household income, from  
all sources, for last year - 1975 - before taxes?"

TABLE 3  
EFFECTS OF INTERVIEWER CHARACTERISTICS UPON  
RESPONSES TO ITEM: "What is your religious  
preference?"

		NA	Refused	DK	Uncertainty Coefficient	N	NA	Prot.	Cath.	Jew.	Other	None	Uncty. Coeff.	N
TOTAL		7.8%	11.0%	27.8%			2.3%	81.7%	4.5%	1.8%	4.5%	5.2%		
Interviewer														
6+ Yrs. Older		4.6	11.1	31.1	.011	901	2.1	77.7	6.5	2.3	4.6	6.8	.023	901
Interviewer														
Within 5 Yrs.		5.0	8.7	29.0		1171	1.0	74.3	4.8	2.0	5.0	13.0		1171
Respondent 5-19														
Yrs. Older		5.7	10.8	25.7		1733	1.7	81.0	5.2	1.9	5.5	4.7		1733
Respondent														
20+ Yrs. Older		9.2	10.9	28.4		3010	1.8	87.3	3.4	1.4	4.1	2.1		3010
Respondent Interviewer														
Male	Female	6.3	11.2	23.7	.003	1272	2.0	73.7	5.6	2.0	5.5	11.2	.010	1272
Male	Male	7.8	12.3	22.9		1590	3.0	81.6	4.8	1.8	3.5	5.3		1590
Female	Female	7.4	8.9	30.0	.002	2137	1.3	82.8	4.0	1.6	5.9	4.4	.006	2137
Female	Male	7.3	11.9	32.7		1949	1.4	87.3	4.0	1.7	3.4	2.1		1949
White	Black	9.6	9.6	32.0	.019	1069	3.2	79.4	5.6	.8	6.7	4.2	.015	1069
White	White	2.8	14.6	19.8		2286	1.4	74.1	7.6	4.8	3.6	8.6		2286
Black	Black	10.5	8.9	31.9	.003	2983	1.7	89.2	1.8	.1	4.4	2.7	.002	2983
Black	White	3.0	7.9	36.2		428	1.6	85.5	2.6	.2	5.1	4.9		428
Means							Means						F	
Interview Length (minutes)		57.42	58.42	57.45			51.40	58.75	55.97	56.35	58.63	55.05	7.01	
Calls Necessary		1.37	1.56	1.34			1.70	1.45	1.57	1.92	1.37	1.69	15.24	
Interview Date		6.45	6.59	6.69			6.46	6.54	6.46	6.21	6.53	6.63	2.38	
Interviewers':														
Age		29.27	33.58	30.57			30.16	31.07	33.27	32.63	29.87	29.30	4.71	
Schooling		15.48	15.78	14.98			15.25	15.31	15.59	15.88	15.18	15.36	3.20	
Days Employed		67.89	70.11	73.12			69.57	67.63	67.93	72.57	73.64	71.91	1.80	
Total Interviews		181.90	180.00	192.60			167.2	178.8	172.4	201.7	198.8	186.8	7.62	
Refusals (number)		36.40	38.44	39.15			36.62	36.51	39.82	50.80	42.91	42.79	19.79	
Training Posttest		60.32	68.70	66.43			68.30	67.20	71.63	78.04	64.60	72.04	16.80	
Evaluation of														
Interviewer		37.31	36.60	37.98			37.09	37.24	38.40	41.47	39.23	38.98	14.38	
Screening Test:														
Total		37.36	37.76	37.56			38.02	37.66	38.02	38.14	37.82	37.90	5.25	
Wording		8.57	8.65	8.56			8.92	8.64	8.78	8.66	8.71	8.66	2.39	
Following Instructions		9.01	9.35	9.26			9.35	9.26	9.57	9.53	9.21	9.44	8.25	
Legibility		9.77	9.83	9.93			9.93	9.87	10.30	9.97	9.85	9.90	2.02	
Audibility		9.93	9.91	10.01			9.80	9.98	9.99	9.98	9.97	9.88	.27	
N		476	522	1680			159	5593	305	121	306	357		