

Effect of Interviewers' Sex on Interview Responses

JOHN COLOMBOTOS, Ph.D., JACK ELINSON, Ph.D., and REGINA LOEWENSTEIN, M.A.

THE HOMOGENEITY of interviewing staffs in sex, age, socioeconomic status, and education has concerned survey researchers for some time. Eighteen years ago, in a summary of the composition of the field staffs of the National Opinion Research Center (NORC), the Gallup and Roper organizations, and other survey research organizations, Sheatsley noted (1):

. . . the composition of most national field staffs has dangerous implications for survey bias arising out of the interviewing situation. We have a condition in which the great bulk of market and opinion research interviewing today is conducted by women talking to men, by college graduates talking to the uneducated, by upper-middle-class individuals talking to those of low socio-economic status, by younger people talking to the increasingly larger old-age groups, by white persons talking to Negroes and by city dwellers talking to rural folk.

Turning specifically to the factor of the sex of interviewers in health surveys, however, the picture is more varied. Of the hundred or so interviewers in the U.S. National Health Survey, all but a handful are women, according to Dr. Philip Lawrence, chief, Division of Health

Statistics, National Center for Health Statistics, Public Health Service.

In two large community surveys on chronic illness, in Hunterdon County, N.J., and Baltimore, Md., all the interviewers were women (2, 3). In other major health surveys, however, including the Kit Carson, Colo., and the Pittsburgh Arsenal Health District studies, the majority of the interviewers were men (4, 5).

One reason for the predominance of women, rather than men, on interviewing staffs is undoubtedly their greater availability for part-time work. A methodological justification is made, however, for the selection of women as interviewers. The justification is based on the assumption that the interviewers' and respondents' being of the same sex facilitates communication between them (6a). The rationale for family health surveys is made explicit by Trussell and Elinson as follows (2a):

Previous studies suggested that the sex of the interviewer has something to do with the yield of reported illnesses. While the data were relatively meager, the evidence was in the direction of supporting the thesis that interviewers of the same sex as the respondent elicited more reports of illness. Since the preferred respondent was the female head of the family, female interviewers were indicated.

But other researchers studying other problems have developed an equally plausible rationale for using male interviewers. For example, Cisin, in a study of drinking practices, noted (7):

Dr. Colombotos is assistant professor, Dr. Elinson is professor and head, Division of Sociomedical Sciences, and Miss Loewenstein is research associate, Division of Medical Care Administration, Columbia University School of Public Health and Administrative Medicine, New York City. This study was supported by the Health Research Council of the City of New York under contract U-1053.

Public opinion poll interviewers are usually middle-aged housewives who look the part. Since male respondents, especially, might be reluctant to discuss the details of their drinking with a lady who subtly reminded them of the village gossip (or, worse yet, of their mothers), the California study made maximum use of male interviewers, selected largely for their studious, academic appearance, an appearance that was supposed to reflect a high level of academic interest and an equally high level of personal disinterest. The manner and attitude of these interviewers were modeled after those of Kinsey's interviewers. Their training reflected the need to give the impression that they had heard everything, that nothing the respondent might reveal could surprise them.

It may very well be, of course, that male interviewers are preferred in certain kinds of health studies, for example, on sexual behavior and alcohol consumption, whereas female interviewers are preferred in other kinds of studies, for example, in the typical family health survey. Sex and liquor are, after all, "man-talk"; the family's health, "woman-talk."

What is the effect of the interviewer's sex when the questions deal with one kind of health data, such as psychiatric symptoms? How would results obtained by an all male interviewing staff differ from results obtained by an all female interviewing staff? An opportunity to answer these questions was provided by the Washington Heights Master Sample Survey, a collaborative project serving several health and medical research groups, at the Columbia University School of Public Health and Administrative Medicine. About two-thirds of the interviewing staff were men. This was not the result of a deliberate policy of recruitment, but rather because interviewers were recruited from college placement offices and the State's employment service.

Our observations are preceded by a brief review of other studies of the effects of interviewers' sex on interview responses, and by descriptions of the master sample survey, the field procedures used, and the two measures of respondents used to compare male and female interviewers.

Review of Other Pertinent Studies

Studies comparing responses obtained by male and female interviewers are meager, possibly because of the small number of men on most interviewing staffs. In a National Opinion Re-

search Center survey in Baltimore in 1947, both male and female respondents gave more puritanical answers to male interviewers than to female interviewers on the following two agree-disagree statements on sexual norms (6b) :

Prison is too good for sex criminals; they should be publicly whipped or worse.

No decent man can respect a woman who has sex relations before marriage.

The effects of the age and sex of interviewers and respondents on the answers to two other questions in the NORC Baltimore survey—one on sex habits and one on a schizophrenic woman—were analyzed by Benney and co-workers (8). In a rather complex interpretation, they concluded that "the least inhibited communication seems to take place between people of the same sex; the most inhibited between people of the same age but different sex." The differences, however, were small and not altogether consistent.

In a Bureau of the Census study, Hanson and Marks found that female interviewers obtained a higher percentage of "females, 14 and older, who are in the labor force" than did male interviewers (9), but it turned out that this difference was because the female interviewers completed more interviews in urban areas, where the proportion of women in the labor force was actually higher.

Turning to health surveys, there was a slight tendency in three studies for male interviewers to obtain more reports of illness and drinking than female interviewers (4a, 5a, 10). The differences, however, were small and limited by the small number of interviewers and the absence of controls on other interviewer characteristics. In one of these studies, moreover, the results are confounded by the fact that all 10 male interviewers were medical students, compared with only two of the eight female interviewers (5a).

Washington Heights Master Sample Survey

The survey data we will examine were collected by interviewers from November 1960 to July 1961. The sample was a two-stage, stratified, clustered sample of 2,300 housing units representing a population of more than a quarter million persons in the Washington Heights Health District in upper Manhattan. Family

forms, in which questions were asked about all members of the household, were completed with 1,942 families. The general format and much of the content of this form were similar to the schedule used by the U.S. National Health Survey. A random sample of about 2,100 adults, 21 years of age and over, was selected from the 1,942 families. From this sample, 1,713 adults were interviewed for about half an hour by interviewers using a respondent form. The questions in this form inquired, in addition to the items examined in this paper, into such diverse topics as the respondent's knowledge, attitudes, and behavior regarding illness and medical care, his integration into groups of friends, his membership in clubs and organizations, and the newspapers he read.

Field procedures. The total sample of 2,300 housing units was divided into three subsamples of equal size. In the second and third subsamples, sets of geographically close housing units were randomly assigned to interviewers. However, because many female interviewers refused to accept assignments in low socioeconomic non-white areas, especially in the southern zone of the district, it was not possible to make assignments that were strictly random throughout the area. Of the randomly assigned interviews completed by men, 35 percent were in the northern zone, 38 percent in the central, and 27 percent in the southern. Of those completed by women, 60 percent were in the northern zone, 36 percent in the central, and only 4 percent in the southern.

These percentages were roughly the same for all respondent forms completed, randomly assigned or not. We decided, therefore, to include all forms in this analysis, whether or not they were randomly assigned, and to control for respondents' race and education, which were strongly associated with zone, in examining differences in the responses obtained by male and female interviewers.

Incidentally, our experience with random assignments of female interviewers points to a dilemma for studies of the effects of interviewers' sex. On the one hand, it indicates that such studies are not normally possible in slum areas, because many female interviewers will not go into these areas. On the other hand, if women are recruited who do agree to accept assign-

ments in these areas, questions may be raised about the degree to which these women are "typical" of female interviewers.

Interviewers in the analysis. Respondent forms were completed by 62 interviewers. The following two controls were imposed on these data: those forms completed by interviewers who interviewed fewer than 15 respondents, and those by the one Negro interviewer who completed more than 15, were omitted from this analysis. This analysis is based on the work of 31 white interviewers who completed 15 or more respondent forms. Their work accounts for 1,479 of the 1,713 forms, or 86 percent.

Characteristics of interviewers. The distributions of the number of respondent forms completed per interviewer were roughly the same among the 21 male and the 10 female interviewers. But the interviewers differed in age, education, and their main activity. Both the men and women were young (only two were over 45), but the men were younger. All but two interviewers, both women, had had at least some college, but the men had had a little more education. More than a third of the men were college or graduate students during the survey. Four of the 10 women were housewives.

Male and female interviewers may differ in other characteristics as well. It is not surprising that survey interviewing, generally an intermittent, part-time job, is likely to attract different types of people from among men and women in the population (1).

The items used. The effects of the interviewer's sex were examined by using two measures of respondents' descriptions of themselves. One measure, called the psychiatric screening score, was based on the number of positive responses to a battery of 22 questions developed by researchers in the Midtown study to distinguish between psychiatrically impaired and well persons (11). These questions were drawn from such sources as the Neuro-psychiatric Screening Adjunct, used in the armed services during World War II, and from the Minnesota Multiphasic Personality Inventory. It included such questions as the following:

Have you ever been bothered by your heart beating hard?

Do you ever have any trouble in getting to sleep or staying asleep?

Have you ever been bothered by cold sweats?
Are you the worrying type?

The second test was based on the number of positive responses to eight questions about anger, guilt, special fears, thoughts and habits, feeling blue or depressed, and feeling very "high," talkative, or active. The questions, suggested by Dr. Bruce P. Dohrenwend and Dr. Lawrence C. Kolb, department of psychiatry, Columbia University College of Physicians and Surgeons, were the following:

Do you have any special fears that keep bothering you?
Do you have any special thoughts that keep bothering you?

Do you have any special habits that keep bothering you?

Do you have periods of feeling blue or depressed which interfere with your daily activities?

Do you have periods of feeling very "high," talkative, or active so that it is difficult to concentrate?

Do you often get angry, irritated, or annoyed?

Do you often feel guilty about things you do or don't do?

Do you often feel frightened or afraid of things?

A positive response to a question indicated a symptom or problem as reported by the respondent. This measure was called the supplementary psychiatric score.

Answers of the Puerto Rican respondents, who make up about 10 percent of the district's population, were excluded from this analysis because they had the highest psychiatric screening scores among the major ethnic groups and because most of them, 102 of 108, were interviewed by men.

Observations

When the respondents' race and sex were controlled in the analysis, male interviewers showed a tendency to obtain higher scores on both the psychiatric screening score and the supplementary psychiatric score from respondents of both sexes and races (see table). The differences between male and female interviewers, however, were small and none were statistically significant. The differences between the average scores obtained by male and female interviewers from all respondents, that is, without controlling for their race and sex, were also negligible: the mean average of the psychiatric screening scores obtained by the male interviewers was 2.17, and by female interviewers, 2.04; the mean average

of the supplementary psychiatric scores obtained by male interviewers was 1.12, and by female interviewers, 0.99.

Although as previously noted, the male interviewers were younger and more likely to be college students than the female interviewers, it was possible that their youth and their status as college students accounted for their obtaining higher scores. The data, when controlled for both age and main activity of interviewer and with Negro respondents excluded because of the small number interviewed by women, did not support this interpretation, however.

The tendency for male interviewers to obtain higher scores from white respondents persisted when age and education of interviewer, sex, age, and education of respondent, and zone were added as controls, but the results were somewhat unstable because of the small numbers of respondents in many cells. Also, this tendency did not appear to be restricted to any group of respondents.

Suppose that an objective of the survey described in this paper had been to examine the relationship between sex of respondents and psychiatric symptoms. Would a staff of male interviewers produce different results from a staff of female interviewers?

The table shows that, regardless of the interviewers' sex, female respondents reported higher scores than male respondents. Furthermore, the differences in both scores between male and female respondents were generally the same whether the interviewers were men or women.

Conclusions

Although there was a slight tendency for male interviewers to obtain more reports of psychiatric symptoms than female interviewers from both male and female respondents, the differences were small and not statistically significant. Essentially, there was no difference.

Whether responses obtained by male and female interviewers differ when respondents report about other members of the family as well as about themselves when asked about acute or chronic physical illness, about visits to the physician, drinking practices, and about other aspects of health and medical behavior, are questions requiring further analysis for a more complete

Psychiatric scores by race and sex of respondent and sex of interviewer

Respondents ¹	Psychiatric screening score				Supplementary psychiatric score			
	Male interviewers		Female interviewers		Male interviewers		Female interviewers	
	Number	Mean average	Number	Mean average	Number	Mean average	Number	Mean average
Total.....	925	2. 17	410	2. 04	884	1. 12	375	0. 99
White:								
Men.....	277	1. 80	148	1. 49	266	. 94	138	. 78
Women.....	365	2. 62	201	2. 44	354	1. 42	183	1. 21
Negro:								
Men.....	113	1. 90	27	1. 81	108	. 84	23	. 78
Women.....	170	1. 99	34	2. 29	156	. 94	31	. 84

¹ Excludes Puerto Rican respondents and respondents not answering all questions for each score.

picture of the effects in health surveys of recruiting male or female interviewers.

Differences in response patterns according to the interviewers' sex may depend on the subject matter and the specific questions asked, in addition, of course, to the respondent populations interviewed and other characteristics specific to a survey. The study results suggest that the rationales commonly presented for hiring male or female interviewers need to be critically re-examined.

Summary

Arguments for recruiting male or female interviewers in health surveys are based on apparently different, though equally plausible rationales. The selection of female interviewers is based on the assumption that similarities between interviewer and respondent, in this case with respect to their sex, facilitates communication between them. Since the respondent in family health surveys is more often the female head of the family, female interviewers are indicated.

In surveys on certain topics, however, such as sexual behavior and drinking practices, it is often argued that male interviewers are more appropriate. The model for this type of survey is the Kinsey interviewer whose qualities were "a high level of academic interest and an equally high level of personal disinterest" that communicated the impression to the respondent that "they had heard everything, that nothing the respondent might reveal could surprise them."

The effects of the interviewers' sex on the re-

spondents' replies were explored during a health survey of a two-stage, stratified, clustered sample of residents of the Washington Heights Health District in upper Manhattan. Respondents answered eight questions about their emotions and 22 questions designed to distinguish between psychiatrically impaired and well persons.

Analysis of results was confined to the responses obtained from 1,479 persons by 31 white interviewers, of which 10 were female.

There was a tendency for male interviewers to obtain higher scores on the two measures used—from both male and female respondents—than female interviewers; however, the differences were not statistically significant.

REFERENCES

- (1) Sheatsley, P. B.: An analysis of interviewer characteristics and their relationships to performance. Pt. 1. *Int J Opinion Attitude Res* 4: 487 (1950-51).
- (2) Trussell, R. E., and Elinson, J.: Chronic illness in a rural area: the Hunterdon study, vol. 3. *In* Chronic illness in the United States, edited by the Commission on Chronic Illness. Harvard University Press, Cambridge, Mass., 1959. (a) p. 61.
- (3) Roberts, D. W., et al.: Chronic illness in a large city: the Baltimore study, vol. 4. *In* Chronic illness in the United States, edited by the Commission on Chronic Illness. Harvard University Press, Cambridge, Mass., 1957.
- (4) U.S. Department of Health, Education, and Welfare: A health study in Kit Carson, Colorado. PHS Publication No. 844. U.S. Government Printing Office, Washington, D.C., 1962. (a) pp. 49, 50.

- (5) Horvitz, D. G. : Sampling and field procedures of the Pittsburgh Morbidity Survey. Public Health Rep 67: 1003-1012, October 1952; (a) p. 1011.
- (6) Hyman, H. H. : Interviewing in social research. University of Chicago Press, Chicago, 1954; (a) pp. 153-170; (b) pp. 165, 166.
- (7) Cisin, I. H. : Community studies of drinking behavior. Ann NY Acad Sci 107: 610 (1963).
- (8) Benney, M., Riesman, D., and Star, S. A. : Age and sex in the interview. Amer J Soc 62: 143-152 (1956).
- (9) Hanson, R. H., and Marks, E. S. : Influence of the interviewer on the accuracy of survey results. J Amer Stat Assoc 53: 635-655 (1958).
- (10) Kirsch, A. D., Newcomb, C. M., and Cisin, I. H. : An experimental study of sensitivity of survey techniques in measuring drinking practices. Social science research project, George Washington University, Washington, D.C., March 1965, pp. 18-20. Mimeographed.
- (11) Langner, T. S. : A 22 item screening score of psychiatric symptoms indicating impairment. J Health Hum Behav 3: 269-276 (1962).

ERRATUM

In the article by Louis Weiner entitled "New York City Department of Health's Weekly Vital Statistics Report" appearing in the May 1968 issue of *Public Health Reports*, pp. 377-382, the charts in figures 1 and 2 were inadvertently reversed. The figures should have appeared as follows:

Figure 1. 1948 death rates per 1,000 from all causes, by weeks, New York City

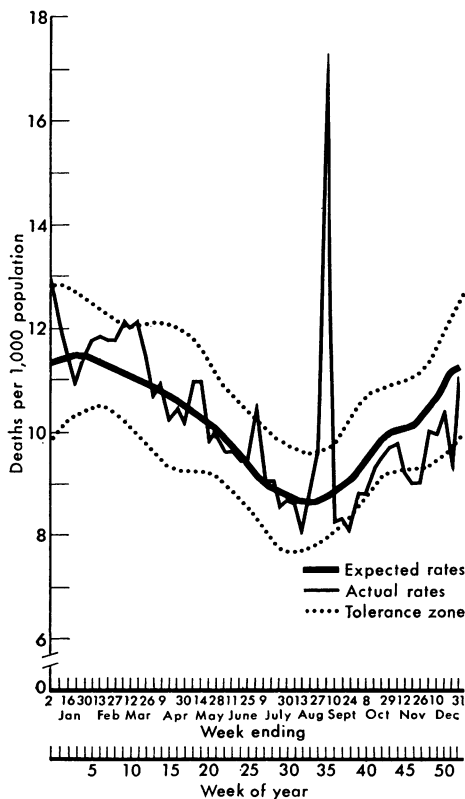


Figure 2. New York City's expected and actual death rates for 1953, by weeks

