INTERVIEWER-RESPONDENT INTERACTION IN A HOUSEHOLD INTERVIEW

Kent H. Marquis, Survey Research Center, The University of Michigan

A recently completed study analyzed the verbal interaction that took place between interviewer and respondent when a revised version of the Urban Employment Survey was used. The technique of verbal interaction analysis is a relatively new approach to studying the household survey. It offers some promising insights.

In 1964, with the help of the National Center of Health Statistics, the U.S. Public Health Service, and the U.S. Bureau of the Census, we set out to test a general model of the dynamics of the interview in which respondent demographic characteristics were hypothesized to cause respondent attitudes, motives, levels of knowledge, and ways of perceiving things. It was thought that these psychological variables, expressed in behavior during the interview, were responsible for the accuracy and completeness of household interview data. Five data collections were involved: 1. Health information from the initial interview with the respondent. 2. Interviewer ratings of each initial interview. 3. Counts of interviewer and respondent behaviors during the interview recorded by a third person. 4. Attitude, motive, knowledge and other data about the respondent collected in an additional interview the day after the original interview. 5. A wide range of data about each of the health interviewers obtained by interviewing them at the end of the study.

The results of this study were quite surprising. Measures of interviewer and respondent attitudes, satisfactions, motivations, knowledge, etc., were not correlated with the amount of health information obtained.[1] On the other hand, the greater the rate of behavioral activity of the respondent, the greater the number of items he reported. This held true after correcting as much as possible for factors causing spurious correlations and regardless of whether the behavior was task-oriented or interpersonally oriented.

This study led us to conclude that the major influences on the accuracy or completeness of data obtained in interviews are the actual behavior of the participants and other parameters which have an effect on the immediate situation (e.g., question characteristics). This conclusion is supported by other programs of research relevant to the interview, such as Matarazzo et al. (2) on interviewee speech duration, and Rosenthal (3) on experimenter bias.

DESIGN AND PROCEDURES

In the design of the following study the respondent's race and age were the main independent variables and the kind of verbal interaction between interviewer and respondent was the dependent variable. Respondent race (black and white) and respondent age (18-34 or 36-64) were combined in a 2 x 2 analysis of variance design. All respondents were employed males. Four female, middle-class, newly trained white interviewers conducted the interviews. Four separate probability samples, one for each of the respondent groups, were drawn from the population of Detroit, Michigan. One-hundred-and-eighty-one usable interviews were obtained, about 45 in each sample group. Several complicating factors in the design and field procedures caused trouble in obtaining satisfactory response rates in inner-city areas. The response rate was between 50 and 60 per cent.

The questionnaire was a modified version of the 1968 Urban Employment Survey of the U.S. Department of Labor. We retained only the questions applicable to the employed and, of these, we discarded many questions that were asked only of sub-groups in the sample.

The verbal interaction throughout the entire interview was recorded with a small, compact, portable tape recorder. The respondent was given the opportunity to refuse to have the interview recorded and one respondent did refuse. Most recordings were coded directly from the tape. We lost seven interviews (4%) because of inaudible recordings.

The system of behavior coding was adapted from the one used in the study of the health interview. We coded reliably observable behavior such as "asks question," "gives response," "probes," "laughs," etc., rather than inferences like "shows hostility," "is enthusiastic," "appears bored," etc. The complete system of codes, summarized in Table 1, contained 36 categories of verbal behavior. At the end of coder training the inter-coder reliability, computed on a code-by-code basis, was between 75 and 80 per cent.

TABLE 1

SUMMARY OF VERBAL BEHAVIOR CODES

Interviewer		Respondent	
Q	Correct question	R	Adequate answer
∠	Incomplete question	W	Inadequate answer
⌐	Inappropriate question	K	Don't know answer
X	Incorrect question	G	Refuses to answer
=	Repeat question	J	Other answer
*	Mistakenly omitted question	E	Elaboration
N	Question omitted because answer already given	C	Asks for clarification
H	Question skipped per instructions		
P	Non-directive probe		
D	Directive probe		
C	Gives clarification		
V	Volunteers information		

Both Interviewer and Respondent

F	Gives feedback
U	Continuing feedback or unsuccessful interruption
T	Repeats answer
A	Irrelevant conversation
S	Gives suggestion
M	Polite behavior
B	Successful interruption
L	Laughs, jokes
Y	Talks to another person
O	Other verbal behavior

RESULTS

Respondents contributed 44 per cent of the verbal behavior in the average interview and interviewers accounted for the remaining 56 per cent. This is in marked contrast to the results of the previous observation study in which respondent behavior constituted a very large majority of the total verbal behavior in the interview.

Interviewer behavior can be classified into three broad areas: question asking, probing, and giving feedback. (See Table 2.) Interviewers are supposed to ask the questions exactly as printed on the questionnaire; according to Table 3, 92 per cent of the questions were asked correctly.

Good interviewer training stresses that non-directive probing is needed. About 22 per cent of all interviewer behavior was devoted to probing of some sort: repeating the question, non-directive probing, directive probing, and repeating answers. For this study, the interviewers behaved the way they were trained. Table 4 shows that about 72 per cent of the probing was non-directive and another 9 per cent consisted of repeating the question, which is another form of non-directive probing. The remaining 19 per cent were directive probes. This was considerably lower than the 42 per cent found in the previous observation study.

TABLE 2

FREQUENCY AND PER CENT OF INTERVIEWER BEHAVIOR, BY CODE CLASSIFICATION

Interviewer Behavior Code Classification	Frequency All Interviews	Per Cent of Total Interview Codes
Asking questions (Q, <, ¬, X)	18,169	37
Probing (=, P, D, T)	11,230	22
Giving feedback (F)	11,498	23
Other (All other codes)	8,948	18
Totals	49,845	100

TABLE 3

FREQUENCY AND PER CENT OF INTERVIEWER QUESTION ASKING BEHAVIOR

Interviewer Behavior	Frequency All Interviews	Per Cent of Total Interview Codes
Questions asked correctly	16,687	92
Incomplete, inappropriate, incorrect questions	1,482	8
Totals	18,169	100

TABLE 4

FREQUENCY AND PER CENT OF 3 TYPES OF PROBING IN 2 STUDIES

Interviewer Probing*	Previous Study Per Cent	Present Study Frequency	Present Study Per Cent
Repeat question	13	638	9
Other non-directive	45	4,938	72
Directive	42	1,269	19
Totals	100	6,845	100

*"Repeats answer" has been left out of the calculations to make the data comparable with the previous SRC observation study.

Table 5 shows that the less directive the probe, the higher the probability of obtaining an adequate answer and the wider the range of behavior which can be elicited. According to theory, non-directive probes should elicit a wide range of behavior. The high rate of adequate answers is less easy to explain. It should be remembered that the interviewer probably does not randomly select the kind of probe that she wants to use and an adequate test of non-directive probing theory requires an experimental design.

The third major classification of interviewer verbal activity is called giving feedback--an interviewer's response to the respondent's behavior ranging from brief signs of approving attention such as "mm-hmm" or "I see" to more elaborate statements like "that's the kind of information we need." Twenty-three per cent of all interviewer behaviors, on the average, fell into this category. This is somewhat disturbing because most programs of interviewer training give very little attention to this aspect of interviewer performance. Two previous studies, which were part of the Survey Research Center Methodological Research Program, anticipated this finding and experimentally varied the interviewer feedback techniques used during the personal interview. They showed that systematic, programmed use of interviewer feedback techniques can increase both the quantity and the accuracy with which health information is reported in the household interview. Data from the present observation study indicate that feedback was used unsystematically. Table 6 shows that the probability that the interviewer gives feedback following an adequate answer is .28.

TABLE 5

PROBABILITIES OF RESPONDENT REACTIONS TO 3 TYPES OF INTERVIEWER PROBES

Repeat Question		Other Non-directive		Directive	
Respondent Reaction	Probability*	Respondent Reaction	Probability*	Respondent Reaction	Probability*
Adequate answer	.38	Adequate answer	.25	Adequate answer	.19
Inadequate answer	.09	Inadequate answer	.08	Other answer#	.65
Other answer#	.05	Other answer#	.42		
Elaboration	.12	Elaboration	.05		
Asks for clarification	.09	Repeats previous answer	.05		
Repeats previous answer	.05				
Talks to third person	.05				

*Reactions with a probability of less than .05 have been omitted.

#Code J, an answer to anything but a correctly asked question (including a probe) which does not meet the objectives of the question on the questionnaire.

TABLE 6

PROBABILITY OF INTERVIEWER FEEDBACK FOLLOWING RESPONDENT BEHAVIOR BY KIND OF RESPONDENT BEHAVIOR

Kind of Respondent Behavior	Probability that Interviewer Feedback Follows
Adequate answer	.28
Inadequate answer	.24
"Don't know" answer	.18
Refusal to answer	.55
Other answer (Code J)	.34
Elaboration	.30
Repeats answer	.32
Gives suggestion	.33
Other behavior (not classified elsewhere)	.21

The probability that feedback is given following an inadequate answer is .24, approximately the same rate. Even more surprising, if the respondent refuses to answer, the probability that the interviewer thanks him or gives other positive feedback is .55. With this kind of feedback schedule, it is understandable that respondents are sometimes confused about what is expected of them. It also indicates a potential source of interviewer bias.

How do respondents behave? The data in Table 7 indicate that 68 per cent of respondent verbal behavior consists of answers to questions and probes. About 14 per cent consists of elaborations upon answers (furnishing information relevant to the general area of the question but which is not responsive to the exact intent of the question). The remaining 18 per cent of respondent behavior is distributed among 10 different verbal behavior categories. It surprised us to find a 6 per cent rate of asking for question clarification (which is high) and a 2 per cent rate of irrelevant conversation (which is very low).

One of the purposes of this study was to find out whether different respondent groups defined

TABLE 7

FREQUENCY AND PER CENT OF RESPONDENT BEHAVIOR, BY CODE CLASSIFICATION

Respondent Behavior Code Classification	Frequency All Interviews	Per Cent of Total Respondent Codes
Answering questions (R, W, K, G, J, T)	26,455	68
Asking clarification (C)	2,431	6
Elaboration (E)	5,520	14
Irrelevant conversation (A)	850	2
Other (All other codes)	3,645	10
Totals	38,901	100

in terms of age and race would show different patterns of interaction with the female, white, middle-class interviewers. To correct for some minor uncontrolled sources of variation, the results are presented in terms of the proportion of total behavior devoted to each code category. Our sample sizes were not quite large enough for many of the between-group differences to be statistically significant. Therefore, trends in the data will be discussed where they appear to present a consistent and meaningful pattern of results.

In general, respondent age accounts for much more variance in verbal behavior than respondent race, with older respondents showing higher frequencies of almost all categories of behavior than younger respondents and with interviewer behavior following this trend. Elsewhere (Cannell et al,1)we have referred to this phenomenon as a "general activity level" effect but have made an attempt to remove it in the present data analysis by the "proportion of total behavior" analysis. Even with this "correction," Table 8 data indicate that interviews with older respondents are characterized by a high percentage of most kinds of behavior. Interviews with younger respondents show higher ratios of asking correct questions, giving adequate answers, and asking for and giving clarifications. The interview with older men tends to be more "diffuse" including higher ratios of probes, inadequate answers, interruptions, elaborations, and irrelevancies. Therefore, the interview with younger men is more to the point or "task-oriented," while greater diversity is characteristic of the interviews with older men.

Behavior patterns differ somewhat between interviews with Negro and white respondents. The major characteristic of interviews with blacks (see Table 9) is the predominant focus on task-oriented behavior (behavior directly relevant to communicating questions and answers). Interviewers devote a higher ratio of their behavior to probing and clarification; black respondents give a higher per cent of answers requiring probing and a higher per cent of requests for clarification.

Initially, it was expected that black respondents would show a lower level of cooperation and young blacks would show some overt resistance. However, the pattern observed was one of cooperative, highly-motivated performance.

White respondents have less difficulty with the task than Negro respondents. Evidence for this includes the finding that interviewers probe and repeat questions proportionately less with white respondents and that white respondents give a higher percentage of adequate answers and fewer inadequate responses.

Interviews with white respondents also show a pattern of greater informality typical of everyday conversation, that is, the verbal exchanges include a greater percentage of non-task behaviors like elaborations and interruptions.

In summary, whites performed adequately and engaged in a higher per cent of informal interaction. Blacks had more difficulty but stuck to the task in a highly motivated, businesslike manner.

CONCLUSION

The coded data contain information about specific problems in the interview, including those indicating incorrectly asked questions, questions which the interviewer omitted by mistake, inadequate answers, requests for clarification, etc. Thus, it is possible to use these data to evaluate problems which occur with the questionnaire. We believe that this kind of approach has considerable potential for use in pretests to provide some objective evaluation of the questions. The coded data also disclose information about interviewer behavior such as her propensity to ask questions correctly or incorrectly, her tendency to use directive or non-directive probes, the frequency with which she accepts inadequate answers to specific questions without probing, etc. The examination of the verbal interaction data can provide us with information for feedback to

TABLE 8

AVERAGE PROPORTIONS OF TOTAL INTERVIEWER BEHAVIOR AND TOTAL RESPONDENT BEHAVIOR BY RESPONDENT AGE

Interviewer Behavior	Average Proportion of Total Older	Younger
YOUNGER > OLDER		
Correct question	.326	.368
Gives clarification	.062	.066
OLDER > YOUNGER		
Incorrect question	.023	.022
Repeats question	.015	.010
Non-directive probe	.106	.095
Directive probe	.027	.022
Repeats answer	.088	.087
Continuing feedback	.028	.018
Interruptions	.003	.002
Irrelevant conversation	.018	.014
Volunteers information	.023	.021
Laughs	.028	.024
OLDER = YOUNGER		
Incomplete question	.006	.006
Inappropriate question	.002	.002
Feedback	.226	.226
Gives suggestion	.002	.002
Polite behavior	.008	.008
Talks to 3rd person	.010	.010
Total Proportions	1.00	1.00
Number of Interviews	90	91

Respondent Behavior	Average Proportion of Total Older	Younger
YOUNGER > OLDER		
Adequate answer	.385	.456
Asks for clarification	.054	.066
OLDER > YOUNGER		
Inadequate answer	.086	.077
"J" answer	.174	.158
Repeats answer	.029	.022
Elaborates	.156	.112
Irrelevant conversation	.022	.016
Continuing feedback	.006	.004
Interruptions	.008	.007
Laughs	.028	.026
Talks to 3rd person	.034	.032
OLDER = YOUNGER		
Don't know answer	.006	.006
Refuses to answer	.000	.000
Polite behavior	.002	.002
Feedback	.007	.007
Gives suggestion	.002	.002
Total Proportions	1.00	.99
Number of Interviews	90	91

interviewers about their performance, to evaluate interviewer training, and to test the effectiveness of the different types of interviewer training. Another area which we would like to pursue would use these kinds of data to study how a pattern of interaction becomes established and what kinds of antecedent conditions predict its appearance. The practical implication of this kind of analysis is the ability to predict in what parts of the interview interaction is smooth, at what point fatigue sets in, and at what point the interaction is developed so that more sensitive issues could be introduced.

We made an initial attempt to use the interaction data to make specific evaluations of questions in the urban employment survey. The results obtained were generally very reasonable and useful. Omission problems appeared mostly in questions with skip patterns or other questions which were asked only of sub-samples. Many of the questions which were asked incorrectly contained ambiguous parenthetical statements or awkward syntax. Inadequate answers from respondents appeared frequently to questions requiring respondent recall of past events or to questions where interviewers were unclear about what constituted an acceptable answer. Questions requiring a large number of behaviors to reach an adequate answer were often the same questions with high inadequate answer rates. With further refinements in the diagnostic approach it should be possible to pinpoint exactly which problem exists for each question, as well as the most likely causes of the problem.

The degree to which the findings of this study can be generalized to all survey interviews is of course limited. There is evidence, for example, that an interview on another topic with different kinds of questions, respondents, and interviewers yields different interactive patterns. The purpose of presenting our data is primarily to demonstrate a methodology which shows promise.

TABLE 9

AVERAGE PROPORTIONS OF TOTAL INTERVIEWER BEHAVIOR AND TOTAL RESPONDENT BEHAVIOR BY RESPONDENT RACE

Interviewer Behavior	Average Proportion of Total White	Negro
WHITE > NEGRO		
Correct question	.358	.334
Volunteers information	.023	.021
Continuing feedback	.025	.022
Interruptions	.003	.002
Laughs	.028	.024
NEGRO > WHITE		
Repeats question	.012	.014
Non-directive probe	.095	.104
Directive probe	.024	.026
Repeats answer	.082	.090
Gives clarification	.062	.065
Feedback	.223	.228
Irrelevant conversation	.014	.018
NEGRO = WHITE		
Incomplete question	.006	.006
Inappropriate question	.002	.002
Incorrect question	.022	.022
Gives suggestion	.002	.002
Polite behavior	.008	.008
Talks to 3rd person	.010	.010
Total Proportions	1.00	1.00
Number of Interviews	90	91

Respondent Behavior	Average Proportion of Total White	Negro
WHITE > NEGRO		
Adequate answer	.428	.414
Refuses to answer	.001	.000
Elaborations	.140	.128
Feedback	.008	.006
Continuing feedback	.006	.004
Interruptions	.008	.007
Irrelevant conversation	.020	.019
Gives suggestion	.002	.001
NEGRO > WHITE		
Inadequate answer	.076	.084
"J" answer	.154	.176
Asks for clarification	.058	.062
Repeats answer	.024	.028
Laughs	.026	.028
NEGRO = WHITE		
Don't know answer	.006	.006
Polite behavior	.002	.002
Total Proportions	1.00	.99
Number of Interviews	90	91

REFERENCES

(1) Cannell, C.F., Fowler, F.J., and Marquis, K.H. "The influence of interviewer and respondent psychological and behavioral variables on the reporting in household interview." National Center for Health Statistics, Public Health Service, PHS Pub. No. 1000, Ser. 2, No. 26, March 1968.

(2) Matarazzo, J.D. "Studies in interview speech behaviors, in L. Krasner and L.P. Ullman, eds. Research in behavior modification. New York: Holt, Rinehart and Winston, 1965.

(3) Rosenthal, R. Experimenter effects in behavioral research. New York: Appleton-Century-Crofts, 1966.

FOOTNOTE

[1]Several earlier validity studies compared respondents' reported hospitalizations, physician visits, or chronic conditions with medical records. These studies revealed a positive correlation between accuracy of reporting and the amount of chronic and acute conditions reported, probably because in health interviews the problem of underreporting is usually greater than the problem of overreporting. Therefore, the first dependent variable of this observation study was an index of chronic and acute conditions reported for the respondent, correcting for gross variations in real health.