

## INTERVIEWER CHARACTERISTICS IN THE CONNECTICUT HYPERTENSION SURVEY

E. Papke, D.H. Freeman, Jr., D.A. D'Atri, A. Ostfeld  
Yale University

### Introduction

The field operations of the Connecticut Hypertension Survey took place between July 20, 1978, and February 16, 1979. During this seven month period, over 100 interviewers were hired and began the training program; of these 72 interviewers engaged in actual field work. Given the large number of interviewers, staffing considerations were critical throughout the survey's field operations. Indeed, these considerations were so critical that the field office staff decided to look carefully at interviewer characteristics and the time, place and manner of interviewer employment and training in hopes of discerning patterns in interviewer productivity. Not only would this study be valuable when the Connecticut Hypertension Survey resumes field operations in the spring of 1982, but it might also contribute to the identification of desirable interviewers in other statewide hypertension studies and in large-scale household surveys in general.

In more specific terms, this paper focuses on both interviewer demographic and conditions of employment variables. In order to control for confounding of these variables, several additional variables relating to the context of the interview are considered. Tests are carried out for association between the assorted variables and interviewer performance rates. An attempt is made to profile the ideal interviewer, and multiple regression analysis is used to test the particular hypothesis that middle-aged women make the best interviewers.

### The Independent Variables

The study gathered data concerning three sets of independent variables: context of the interview, interviewer demographics and conditions of employment. The first set, context of the interview, was gathered from the final computerized interviewer control report. These data, calculated by interviewer, included the respondents' mean age, respondents' mean systolic and diastolic blood pressures (using the second and third readings), proportions of six sex-race combinations of respondents, mean visit on which enumeration was obtained, and mean visit on which the interview was obtained. The second set, demographic data regarding the interviewers, was gathered from personnel applications, time sheets and field office records. These data included age, sex, race, marital status, and number of years of school completed.

Finally, data on conditions of employment were gathered from the same sources and included training session, office assignment, prior interview experience, attendance at a group meeting with the principal investigators, listing experience, length of time in the field interviewing, and reason for leaving the job. While the context of the interview and interviewer demographic variables are self-explanatory, the conditions of employment variables are more unusual and reflect the particular conditions of the Connecticut Hypertension Survey. In preparation for and in the actual course of field work, field office staff made numerous decisions affecting conditions of employment. To cite only two examples, field staff

found it necessary to institute eleven training cycles and to establish a satellite office in Hartford.<sup>1</sup> Since the variables derived from these administrative decisions are specific to the Connecticut Hypertension Survey, they are of course less directly applicable to other survey experiences than are the context of the interview and interviewer demographic variables.

### The Dependent Variables

Interviewer performance rates, the dependent variables, were defined as the enumeration rate, interview rate, net response rate and production rate. The enumeration and interview rates were taken from the final computerized interviewer control report. The enumeration rate measures the interviewer's success in obtaining initial information regarding the residents of the household; the rate is in essence a measure of the interviewer's aggressiveness in getting in the door. The interview rate, meanwhile, is a measure of the interviewer's ability to complete the personal interview which included three blood pressure readings. The net response rate in turn is the product of the enumeration and interview rates. Finally, the interviewer production rate is the number of interviews completed divided by the number of weeks in the field. These rates apply only to Project 003 of the survey, the actual data collection period. In other words, interviewer production and response rates in the reliability study, or Project 004, have not as yet been addressed.

### Methods of Analysis

Sixty-nine interviewers are included in the data analysis. Of the 76 interviewers who completed training, four were excluded because they did not complete any interviews, and three were excluded because there was information missing. Data analysis consisted of two steps and used the Statistical Analysis System (SAS). Descriptive statistics on interviewer demographic and conditions of employment variables were obtained in the first step and will be presented shortly. The descriptive statistics for each of the continuous variables are presented in Table 1. Multiple regression analysis was then performed to determine the association between demographic and conditions of employment variables and interviewer success rates, controlling for the context of the interview. The four dependent variables--enumeration rate, interview rate, net response rate and production rate--were each examined with respect to the independent variables outlined above. The General Linear Models procedure (GLM) was applied to all independent variables (see Table 2). Those variables which were not significant at  $p < .25$  were eliminated from each of the four models. Then GLM was run on the remaining variables. In the next step, variables which were not significant at  $p < .10$  were eliminated. The final models include those variables significant at  $p < .05$  and are presented in Table 3.

### Results

The first step of analysis summarizes interviewer demographics and conditions of employment. Of the sixty-nine interviewers included in the analysis, 49 or 71% were female and 20 or 29% were

male. Sixty-three or 91% were white while 6 or 9% were members of racial and ethnic minorities. Forty-seven or 63% of the interviewers were single. Thirty-one interviewers or 45% had a college education; 14 or 20% had completed one or two years of school beyond college. The interviewers ranged in age from 17 to 57 years old. With respect to conditions of employment, 7 interviewers or 10% had prior interviewing experience. Thirty-seven interviewers were based at the New Haven office while 32 were based in Hartford. Thirty-nine interviewers attended a group meeting with the principal investigators while 30 did not. Fifteen interviewers or 22% left the survey at the end of summer, 1978; this group consisted of college students hired specifically for the summer. Eleven or 16% left to take another job. Ten or 14% left because they did not like the job. Fourteen or 20% stayed until the end of field work in February, 1979. Finally, 19 interviewers or 28% were terminated after discussion with the field director. In short, the Connecticut Hypertension Survey recruited and trained interviewers with a wide range of demographic characteristics. These interviewers in turn had a wide range of employment experiences.

Before turning to the results of the GLM procedure, it should be noted that interviewer assignments were not randomized. Therefore, the staff controlled for the context of the interview in order to develop more reliable regression models. Further study would look more carefully at the interaction of interviewer and respondent characteristics based on randomized assignments.

#### (1) Results Regarding Demographic Characteristics

The final GLM models reveal that the standard demographic variables are not significant determinants of interviewer performance rates. The variables age and sex are not significant at  $p < .25$  in any of the four models. Of the remaining three variables--marital status, race and number of years of school completed--marital status is not significant at  $p < .10$  in the model for net response rate; race is not significant at  $p < .10$  in the model for production rate; and number of years of school completed is not significant at  $p < .10$  in the model for interview rate.

These data, therefore, are not conclusive with respect to the hypothesis that middle-aged women make the best interviewers. Field office staff members were particularly interested in determining if this were the case. Indeed, the data indicate that there are no significant differences among female and male interviewers.<sup>2</sup> Field office staff originally considered male applicants less desirable as interviewers due to a fear that people would be hesitant to open their doors to them. Given these data, it seems that this concern need not be a factor in the consideration of specific interviewer attributes during the hiring process.

#### (2) Results Regarding Conditions of Employment

Those variables which are in fact significant predictors are not demographic in nature but rather related to the conditions of employment. The first is the training session variable, and as the summary of the models in Table 3 indicates, training session is significant in three of four models. The p values for interview, net response

and production rates are, respectively, .0012, .0032, and .0002. On reflection, this finding is not surprising since the training session is the critical mechanism for communicating interview techniques and skills; in the end mastery of these techniques and skills influences interview and overall response rates. Particularly noteworthy is the central placement and importance of three blood pressure readings in the interview. Each reading constituted personal, physical contact with the respondent, and it was greatly advantageous if the interviewer could manifest confidence and expertise in the use of the sphygmomanometer. Apparently, the various training sessions more or less successfully developed these attitudes and skills. However, the training process deserves additional scrutiny before any specific recommendations can be made.

A second variable related to conditions of employment which is significant is office assignment. As Table 3 illustrates, office assignment is significant in the model of interview rate at  $p = .0097$  and in the model of net response rate at  $p = .0067$ . These models suggest that overall structure and support mechanisms of the field operations are important factors to consider with respect to productivity and response rates. Moreover, the mean enumeration, interview, net response and production rates obtained by the two offices are significantly higher for interviewers based at the Hartford office than those based in New Haven ( $p = .05$ ). Reporting to one coordinator rather than the main field office staff simplifies the office procedures from the interviewer's perspective. As the satellite office, Hartford was removed from the intensity of activity in New Haven. These findings seem to confirm the belief that a decentralized field operation would improve overall survey efficiency even in a state as small as Connecticut.<sup>3</sup>

#### Conclusion

These data suggest that it is not possible to present a demographic profile of the ideal interviewer. On the contrary, such conditions of employment variables as training session and office assignment are more significant determinants of interviewer performance than the demographic characteristics of interviewers. Although the sample of interviewers from the Connecticut Hypertension Survey is small, one tentatively concludes that the management and direction of the survey are most important to high interviewer productivity. Particularly critical are the hiring and training processes, maintenance of morale among interviewers and survey organization. A well planned, directed and managed survey as opposed to a demographically distinct interviewing team seems the key to high interviewer performance in a large-scale household survey.

#### References

1. For a complete discussion of field operations of the Connecticut Hypertension Survey see D.A. D'Atri et al., "Plan and Operation of the Connecticut Hypertension Survey," American Statistical Association 1979 Proceedings of the Section on Survey Research Methods, pp. 179-181.
2. John Freeman and Edgar W. Butler, "Some Sources of Interviewer Variance in Surveys," Public Opinion Quarterly, 40:1, 79-91, 1976. In a small

study which examined interviewer variance, Freeman and Butler found significant variance between male and female interviewers. Sex of respondent, however, was not considered.  
3. D.A. D'Atri, op. cit., p. 180.

### Acknowledgement

The authors wish to thank K. Hellenbrand for her help. This work was supported in part by the Connecticut Health Department, Contract Number 79-131 and the National Heart, Lung and Blood Institute, Contract Number N01H-72984.

Table 1

Descriptive Statistics for Interviewer Conditions of Employment Variables and Characteristics of Each Interviewer's Respondents Connecticut Hypertension Survey

Variable	N	Mean	Std. Dev.	Skewness
Length of time interviewing (weeks)	69	5.338	3.535	1.171
Interview rate	69	.811	.131	-.673
Enumeration rate	69	.928	.091	-1.442
Net response rate	69	.757	.160	-.486
Production rate	69	11.695	5.650	.868
Meanage	69	43.692	4.819	-1.225
Prblkmn	69	.025	.045	2.499
Prblkwom	69	.046	.128	6.389
Prwhtmen	69	.383	.134	-1.438
Prwhtwom	69	.496	.160	.755
Prothmen	69	.026	.045	3.598
Prothwom	69	.023	.043	4.148
Meansys	69	126.157	6.223	-.781
Meandias	69	78.957	3.162	-.506
Intemean	69	1.429	.322	1.828
Enummean	69	2.131	.728	.764

Table 2

Estimates for Interviewer Demographic, Conditions of Employment and Characteristics of each Interviewer's Respondents Parameters Connecticut Hypertension Survey

Parameter	Enumeration Rate Estimates	Interview Rate Estimates	New Response Rate Estimates	Production Rate Estimates
Intercept	.998*	.478	.638	28.042
Age	.000	.001	.001	-.010
Sex <sup>a</sup>	-.001	-.013	-.012	1.830
Marstat <sup>b</sup>	-.027	-.030	-.049	1.617
Race <sup>c</sup>	.042	-.043	.001	7.326*
Schyrs	-.003	.014	.011	.044
Listed <sup>d</sup>	-.069	-.105	-.116	4.199
Intexp <sup>e</sup>	.057	-.016	.038	-1.340
Lotint	-.007	-.007	-.011	.300
Training Session <sup>f</sup> 1	-.056	.056	.031	10.107
2	-.002	.165*	.152	8.319*
3	-.045	.169*	.118	2.721
4	.081	.224**	.266*	9.816*
5	-.005	.166*	.145	3.191
6	-.009	-.033	-.034	-.177
7	-.067	.071	.022	-1.447
8	.009	.118	.102	.673
9	.022	.110	.117	7.633
10	.020	.106	.119	4.346
Office <sup>g</sup>	-.014	-.106**	-.104*	-2.682
Printg <sup>h</sup>	.048	-.025	.015	.700
Leave <sup>i</sup> 1	-.019	.018	.003	-4.104
2	-.030	-.056	-.074	2.418
4	.030	.038	.045	-3.796
5	.021	-.094*	-.075	2.370

Table 2

Estimates for Interviewer Demographic, Conditions of Employment and Characteristics of each Interviewer's Respondents Parameters Connecticut Hypertension Survey

Parameter	Enumeration Rate Estimates	Interview Rate Estimates	New Response Rate Estimates	Production Rate Estimates
Meanage	.003	-.006	-.003	.155
Prblkmn <sup>j</sup>	-.036	.532	.297	-2.165
Prblkwom <sup>j</sup>	.036	1.616**	1.370*	-21.081
Prwhtmen <sup>j</sup>	.031	.998*	.770	-19.257
Prwhtwom <sup>j</sup>	.000	1.111*	.882	-33.215
Prothmen <sup>j</sup>	-.724	1.239	.150	-69.884
Meansys	-.005	.004	.000	.051
Meandias	.007	-.009	-.003	-.151
Intemean	.003	-.206**	-.187*	-2.257
Enummean	-.053**	-.036	-.073**	-1.711

- a. relative to male
- b. relative to not married/never married
- c. relative to racial/ethnic minorities
- d. relative to those with listing experience
- e. relative to those with prior interviewing experience
- f. relative to training session 11.
- g. relative to Hartford Office
- h. relative to those who attended meeting with principal investigators
- i. relative to reason 9 - left, mutual agreement; fired
- j. relative to proportion of other women

\* Significant at p < .05

\*\* Significant at p < .01

Table 3

Final Regression Models  
for Interviewer Characteristics in the  
Connecticut Hypertension Survey (continued)

A. Enumeration Rate

ANOVA Table				
Source	df	Sum of squares	F value	p value
Model	3	.182	10.37	.0001
Error	65	.381		
Total	68	.563		
Meansys	1	.044	7.54	.0078
Meandias	1	.067	11.35	.0013
Enummean	1	.125	21.30	.0001

Parameter Estimates

Parameter	estimate	t test	p value	Standard error
Intercept	.758	2.93	.0046	.258
Meansys	-.004	-2.75	.0078	.002
Meandias	.011	3.37	.0013	.003
Enummean	-.060	-4.61	.0001	.013

Table 3

Final Regression Models  
for Interviewer Characteristics in the  
Connecticut Hypertension Survey (continued)

C. Net Response Rate

ANOVA Table				
Source	df	Sum of Squares	F value	p value
Model	16	1.324	10.15	.0001
Error	52	.424		
Total	68	1.748		
Training session	10	.257	3.15	.0032
Office Assignment	1	.065	7.98	.0067
Prblkwom	1	.249	30.57	.0001
Prwhomen	1	.096	11.75	.0012
Prwhwom	1	.181	22.15	.0001
Intemean	1	.075	9.19	.0038
Enummean	1	.175	21.46	.0001

Parameter Estimates

Parameter	Estimate	t test	p value	Standard error
Intercept	.658	5.12	.0001	.128
Training Session	1	.113	2.19	.0332
	2	.165	2.98	.0044
	3	.043	.79	.4351
	4	.211	3.33	.0016
	5	.094	1.75	.0868
	6	-.048	-.82	.4144
	7	-.026	-.41	.6802
	8	.040	.67	.5054
	9	.045	.82	.4150
	10	.103	2.05	.0457
Office Assignment	-.101	-2.82	.0067	.036
Prblkwom	1.123	5.53	.0001	.203
Prwhomen	.493	3.43	.0012	.144
Intemean	-.172	-3.03	.0038	.057
Enummean	-.097	-4.63	.0001	.021

Table 3

Final Regression Models  
for Interviewer Characteristics in the  
Connecticut Hypertension Survey (continued)

B. Interview Rate

ANOVA Table				
Source	df	Sum of Squares	F value	p value
Model	14	.819	9.24	.0001
Error	54	.342		
Total	68	1.161		
Training Session	10	.225	3.55	.0012
Office Assignment	1	.046	7.19	.0097
Proportion				
Black women				
Interviewed	1	.119	18.87	.0001
Intemean	1	.080	12.64	.0008
Enummean	1	.147	23.24	.0001

Parameter estimates

Parameter	estimate	t test	p value	Standard error
Intercept	1.173	17.18	.0001	.068
Training session	1	.091	2.05	.0448
	2	.151	3.17	.0025
	3	.024	.53	.5966
	4	.164	2.97	.0045
	5	.092	1.97	.0538
	6	-.080	-1.57	.1212
	7	.005	.09	.9258
	8	.045	.84	.4028
	9	.042	.91	.3692
	10	.096	2.20	.0322
Office Assignment	-.084	-2.68	.0097	.031
Prblkwom	.446	4.34	.0001	.103
Intemean	-.156	-3.56	.0008	.044
Enummean	.083	4.82	.0001	.017

Table 3

Final Regression Models  
for Interviewer Characteristics in the  
Connecticut Hypertension Survey (continued)

D. Production Rate

ANOVA Table				
Source	df	Sum of Squares	F value	p value
Model	11	995.208	4.39	.0001
Error	57	1175.650		
Total	68	2170.858		
Training Session	10	889.172	4.31	.0002
Prothmen	1	185.754	9.01	.0040

Parameter Estimates

Parameter	Estimate	t test	p value	Standard error
Intercept	7.681	4.46	.0001	1.721
Training Session	1	3.734	.0823	2.111
	2	7.620	2.87	.0058
	3	9.678	3.83	.0003
	4	10.519	3.67	.0005
	5	6.605	2.58	.0125
	6	1.824	.63	.5288
	7	.638	.20	.8395
	8	1.828	.69	.4948
	9	10.981	4.35	.0001
	10	3.119	1.28	.2041
Prothmen	-40.200	-3.00	.0040	13.395