EXPERIMENTS IN TELEPHONE-PERSONAL VISIT SURVEYS

John M. Bushery, Charles D. Cowan and Linda R. Murphy
U. S. Bureau of the Census

Traditionally, personal visit interviewing has been the method preferred by the Census Bureau for conducting demographic sample surveys. The consensus of opinion has been that, in a face-to-face situation, the interviewer can more easily establish his identity and the legitamacy of the survey, allay respondent suspicions, and avoid refusals and terminations in midinterview, thus producing higher response rates. Further, the general belief is that personal visit interviewing results in better quality data because the interview can be more thorough, and because problems in communication are reduced when the interviewer can use visual cues to help the respondent with difficult portions of the interveiw. Finally, it is felt that respondents tend to prefer the face-to-face interview experience. A paper by John Scott has already discussed this problem [8].

Unfortunately, personal visit interviewing is already very expensive and becoming more expensive all the time. So sponsors of surveys are investigating the use of less expensive alternatives to personal visit interviewing. One of the more promising alternatives appears to be telephone interviewing.

Many of the surveys carried out by the Bureau of the Census require long and detailed interviews. Recently the Census Bureau conducted controlled experiments in two major ongoing surveys--the National Crime Survey (NCS, sponsored by the Law Enforcement Assistance Administration) and the Current Medicare Survey (CMS, sponsored by the Social Security Administration). The purpose of these experiments was to carefully examine the differences between data collected primarily by telephone and data collected primarily in person. The two experiments had several similar features. Each survey involved repeated interviewing of respondents over time. For each experiment, subsamples were designated to be interviewed using procedures which maximized either personal visit or telephone interviewing. These experimental procedures were tested within the constraints of the existing Census Bureau field organization. This involved "preconditioning" each household or respondent with at least one personal visit interview. The experimental procedures actually involved a mixture of personal visit and telephone interviewing. For example, telephone interviews were allowed in the personal visit treatment, but only as a last resort, to avoid noninterviews. Further, in each experiment certain households (or persons) were administratively ruled to be ineligible for a telephone interview. Only those households (or persons) actually interviewed in the previcus enumeration could be interviewed by telephone. Those not interviewed in the previous enumeration were to be interviewed in person. These ineligible cases consisted of noninterviews in the previous enumeration, cases added to update the sample, and (in the NCS) households which had moved into a sample unit since the previous enumeration. Telephone interviews were not allowed in these cases because it was felt that face-to-face contact was necessary to introduce new respondents to the survey and to re-establish a rapport between interviewer and respondent in cases of a previous noninterview.

Despite these similarities, the two experiments had many dissimilar features. The most important of these is the difference between the designs of the two experiments. For this reason, the discussions of the experimental designs, the analyses performed, and the primary results of the two experiments will be discussed separately. The conclusions based on these experiments will, however, be discussed jointly.

The NCS Experiment--Design

The NCS experiment was conducted in an effort to determine whether a procedure which maximized personal visit interviewing or a procedure which maximized telephone interviewing would produce estimates different from those produced by the current NCS interview procedure. The current procedure involves mostly personal visit interviewing, with some telephone interviewing for callbacks. Two systematic subsamples, referred to as the personal visit group and the telephone group, were selected from the NCS sample and interviewed using the experimental interview procedures. Each subsample, consisted of about 6,000 households and accounted for approximately one-twelfth of the NCS sample. The remaining five-sixths (about 60,000 households) were interviewed using the current NCS interview procedure and served as a control in the experiment.

The current NCS procedure involves the use of both a household respondent and self-response by each household member 14 or more years old.[1/] The household respondent, an adult member of the household, must be interviewed in a personal visit. Other available household members are interviewed at the same time. Callbacks to absent members of the household may be conducted by telephone or in person, as the interviewer chooses. This procedure results in face-to-face interviews with about 78 percent of all respondents. (See table 1.)

As mentioned earlier, telephone interviews were permitted in the personal visit group only as a last resort, to avoid noninterviews. This resulted in personal visit interviews with about 96 percent of the respondents in the personal visit group. However, only 88.4 percent of the households in the telephone group were eligible to be interviewed by telephone. These were the "identical repeat" households which had actually been interviewed during the previous enumeration. The ineligible 11.6 percent were interviewed using a procedure identical with the current NCS procedure, except that all callbacks were to be made by telephone, if at all possible. This resulted in telephone interviews with about 80 percent of the respondents in the telephone group. (See table 1.)

The NCS Experiment--Analyses Performed

In the NCS experiment, data were collected from July 1976 through June 1977 and two similar, but distinct analyses were carried out.

The first analysis [5] was primarily concerned with the differences that could be expected if either experimental interview procedure were used to replace the current NCS procedure. To accomplish this, the victimization rates for the control group were compared with the corresponding rates for the personal visit and telephone groups. No direct comparisons were made between the two experimental groups.

In the second analysis, however, the levels of reporting obtained using personal visit and telephone interviewing were compared directly. Only data collected from "identical repeat" households were used in this analysis. Unfortunately, the variances on the estimated victimization rates in the two experimental groups were fairly large. This made it difficult to find significant differences between the two experimental procedures. Nonparametric methods were therefore used in an effort to detect whether any fundamental differences exist between the two procedures. The findings of this analysis are discussed in [4].

In both analyses the assumption has been made that higher reported victimization rates represent "better" data. This assumption appears reasonable since early research in the NCS indicated that victimizations tend to be underreported [6].

The NCS Experiment--Findings

Limitations of space permit only a brief summary of the major findings reported in [5], some of which are illustrated in table 1.

In order to determine whether interviewing procedure had an effect on nonresponse rates, the noninterview rates for the two experimental groups were compared with those for the control group. In the NCS experiment two types of noninterview were of particular interest, type A and type Z noninterviews.[2/] There was no evidence that the type A noninterview rate is affected by interview procedure. (See table 1.) On the other hand, the type Z non-interview rate in the telephone group, 2.4 percent, was significantly higher than the type Z rate in the control group, 1.9 percent. Although this difference is statistically significant, it is quite small and should have only a negligible effect on data quality.

Although there is no completely consistent pattern, the data suggest that a higher proportion of personal visit interviews tends to result in higher victimization rates. That is, victimization rates in the personal visit group were usually at about the same level as, and sometimes higher than the corresponding rates in the control group. Victimization rates in the telephone group, on the other hand, were usually lower than or at about the same level as the corresponding rates the the control group.

For total crimes against persons, crimes of violence and crimes of theft, there is no evidence that the reporting rates in the control group and the personal visit group were different. However, the rate for total crimes against persons in the telephone group, 119.1 crimes per thousand persons, is significantly lower than the rate for the control group, 129.4 crimes per thousand. The primary reason for this difference appears to be a lower reported rate for crimes of theft without contact. In particular, underreporting of these crimes by whites and males appears to contribute most to the lower overall victimization rate in the telephone group. Closer examination has revealed that almost all the difference between the control group in the rates for crimes of theft without contact occurs among thefts of property with a reported value of less than $25. (See [5].)

Other important differences can be seen in the victimization rates for aggravated assault reported by blacks. Blacks in the personal visit group reported total aggravated assaults and aggravated assaults committed by nonstrangers at significantly higher rates than their counterparts in the control group. In addition, blacks in the telephone group reported aggravated assaults committed by strangers at a significantly lower rate than their counterparts in the control group. (See table 1.)

The findings in [4] suggest that respondents in "identical repeat" households tend to report crimes at a higher rate when interviewed in person than when interviewed by telephone.

Finally, choice of interview procedure may affect comparisons between population subgroups. For example, in the telephone group, both blacks and whites reported crimes of violence at about the same rate. In the control group, however, blacks reported crimes of violence at a rate about 38 percent higher than the rate for whites; and in the personal visit group the victimization rate for blacks was 77 percent higher than the rate for whites. (See table 1.) This example involves differences among the three interviewing procedures which are not statistically significant. Nevertheless, it serves to illustrate the substantial effect choice of interview procedure may have on comparisons between population subgroups.

The CMS Experiment--Design

The purpose of the Current Medicare Survey was to obtain estimates of the frequency of use for various forms of medical care. The variables of interest include visits to doctors, visits to hospitals, use of medical services, stays in nursing homes, and use of prescriptions. The CMS also collected information on Medicare coverage of expenditures for these items. The CMS sample of about 6,800 individuals was selected from the Social Security Administration files of aged and disabled persons. Individuals in sample were

interviewed for 15 successive months, starting in October of each year and continuing through December of the following year.

The original CMS procedure was to interview all sample individuals in person. However, in June 1974 telephone interviewing was introduced to the survey, with telephone interviews used in two of every three months. Personal visit interviews were used every third month because it was felt that periodic face-to-face contact with the interviewer would help keep response rates high. There was also some concern that data quality might deteriorate if telephone interviewing were used exclusively over a period of several months.

A preliminary study conducted in 1974 [3] found no appreciable differences between telephone and personal visit interviewing. However, the samples used were rather small, so a larger experiment was carried out over eight months, from May through December, 1976.

In this experiment the maximum personal visit and the maximum telephone treatments were compared directly. No comparisons were made with a "standard" or control treatment.

In order to assign interview treatments, primary sampling units (PSU's) were grouped into homogeneous clusters of size three and each PSU in a cluster was assigned to one of three randomization groups. The interview treatment to be used for each group in any given month was specified in the design matrix shown in Figure 1. This design was used because it retained the feature that every third interview would be conducted face-to-face.

Figure 1. Assignment of Interview Treatment for PSU's in the CMS Experiment

Month of Interview	Randomization Group		
	A	B	C
May 1976	P	T	T
June 1976	T	P	T
July 1976	T	T	P
August 1976	P	T	T
September 1976	T	P	T
October 1976	T	T	P
November 1976	P	T	T
December 1976	T	P	T

T = Telephone treatment P = Personal Visit treatment

The administrative rules governing the use of personal visit and telephone interviews resulted in the use of face-to-face interviews about 88 percent of the time in the personal visit treatment, and telephone interviews about 58 percent of the time in the telephone treatment. It is suspected that the relatively low porportion of telephone interviews conducted in the telephone treatment may be due to the restricted target population in the CMS, aged and disabled persons.

The CMS Experiment--Analyses Performed

The primary analysis in the CMS experiment consisted of testing for differences between estimated mean levels of reporting for the eight month period. In the telephone treatment these estimates were based on about 36,000 interviews and in the personal visit treatment on about 18,000 interviews. Since validation efforts in the health interviewing area have revealed that higher levels of reporting need not represent "better" data, no attempt has been made here to choose either treatment as producing more accurate data.

The t-test was used to detect differences between the two treatments. However, tests for heteroscedasticity have revealed that the variances obtained from the two treatments are usually different. This represents a classic case of the Behrens-Fisher problem, with unequal treatment variances. An adjustment to the degrees of greedom in the t-test as described in [2] has been made to correct for the effects of heteroscedasticity. In addition, the data yielding the telephone and personal visit estimates were not independent. A preliminary analysis of the correlations involved suggests that the t-test used here is more liberal than would be a t-test on independent data. As a result, this analysis is more likely to falsely conclude that differences exist between the mean levels of reporting in the two treatments. However, as will be seen, no such differences were found, even using this liberal t-test.

The CMS Experiment--Findings

There is no evidence that interview procedure has any effect on the nonresponse rate. The data tapes provided by the Social Security Administration contain imputed data for all noninterview cases. One would expect 66.7 percent of all imputed cases to occur in the telephone treatment, if chance alone were operating. The actual proportion, 66.5 percent, of imputed cases in the telephone treatment is not significantly different from the expected proportion. For the 21 variables shown in table 3, there were no significant differences by interview treatment in the mean levels of reporting. This corroborates the 1974 study.

As mentioned earlier, tests for heteroscedasticity were conducted. These tests formed another part of the CMS analysis. For all but two of the variables in table 3, the two interview treatments resulted in significantly different variances. In most cases the variance for the personal visit treatment was larger. Fortunately, none of these differences was large enough to have any practical effect on the relative reliability of the means estimated, given the sample sizes used in the CMS.

Finally, a two-way analysis of variance was performed for seven variables: Number of visits to doctors, Number of times in hospital, Number of times in nursing homes, Number of ambulance uses, Number of days in bed, Total number of items not covered by Medicare, and Total charges for items not covered by Medicare. The purpose of this analysis was to verify that the design of the experiment prevented a divergence of telephone and personal visit estimates over time. For only

one variable, Number of times in a nursing home, was there evidence of an interaction between treatment and month.

Conclusions

The results of the NCS experiment indicate that personal visit interviews tend to produce slightly better victimization data than the current NCS procedure. The current procedure, in turn, tends to produce slightly better data than telephone interviewing. It also appears possible that a poor choice of interviewing procedure may introduce some bias into comparisons between population subgroups. Despite the fact that telephone interviewing does not produce the highest quality data in the NCS, its deficiencies do not appear to justify outright rejection of a maximum telephone procedure for use at some future time. It may be possible to develop techniques which compensate for the weaknesses of telephone interviewing. This could permit telephone interviewing to achieve results comparable with those of personal visit interviewing.

Since a large portion of interviewing costs involves travel time and reimbursement for mileage, the savings of telephone interviewing over the current NCS procedure could be substantial. Rough estimates indicate that the average interviewing cost, per case, in the telephone group was only three-fourths as high as the cost in the control group [7].

The major conclusion to be drawn from the CMS experiment and from a Census Bureau experiment conducted in the Household Survey of Residential Alteration and Repairs (SORAR, see reference [1]) is more encouraging: apparently telephone and personal visit interviewing produce comparable data.

It should be remembered that in these experiments all the interview treatments actually involved a mixture of interview modes. That is, the maximum telephone treatment included a sizeable proportion of personal visit interviews and the personal visit treatment included at least a few telephone interviews.

These studies, taken together, demonstrate that the effectiveness of telephone interviewing depends in large part on the information required from the respondent. Thus, for surveys similar to the CMS, telephone interviewing would probably produce data consistent with personal interviewing, at substantial savings. However, caution must be exercised. Telephone interviewing may not always be a viable alternative, particularly if the interview is long and complicated. Adequate testing should be performed prior to introducing telephone interviewing into any survey.

Footnotes

1/ Interviews for 12 and 13 year olds are obtained in a proxy interview.

2/ A type A noninterview occurs when the interviewer fails to conduct an interview at an occupied housing unit or other type of living quarters. A type Z noninterview occurs when no interview is obtained for a person within an interviewed household.

References

[1] Champion, Elinor., "Update of Use of Telephone Interviewing in SORAR." Washington, D. C.: U. S. Bureau of the Census, Construction Statistics Division, April 5, 1978.

[2] Hays, William L., Statistics for Psychologists. New York: Holt, Rinehart, and Winston, 1963.

[3] Koons, David A., "Current Medicare Survey, Telephone Interviewing Compared with Personal Interviews," Response Research Staff Report #74-4. Washington, D. C.: U. S. Bureau of the Census, Statistical Research Division, October 4, 1974.

[4] Turner, Anthony G., "An Experiment to Compare Three Interviewing Procedures in the National Crime Survey." Washington, D. C.: U. S. Bureau of the Census, Statistical Research Division, December 1977.

[5] Woltman, Henry F. and Bushery, John M., "Results of the NCS Maximum Personal Visit-Maximum Telephone Interview Experiment." Washington, D. C.: U. S. Bureau of the Census, Statistical Methods Division, December 9, 1977.

[6] San Jose Methods Test of Known Crime Victims (Statistical Technical Report No. 1) United States Department of Justice, LEAA, National Institute of Law Enforcement and Criminal Justice, Statistics Division, June 1972, Washington, D. C.

[7] Memorandum for Richard F. Blass from Mary Ellen Beach and Marcia Thessin; "Comparison of NCS Interviewing Costs Derived from Maximum Telephone/Maximum Personal Visit Experiment"; Washington, D. C.: U. S. Bureau of the Census, Field Division, June 14, 1978.

[8] Scott, John C., "Attitudes toward Survey Modes" University of Michigan, paper presented at 1978 Meetings of American Statistical Association.

TABLE 1. SOME RESULTS OF THE NCS MAXIMUM PERSONAL VISIT-MAXIMUM TELEPHONE INTERVIEW EXPERIMENT

Characteristic	Procedure			Standard versus Personal Visit		Standard versus Telephone	
	Standard	Personal Visit	Telephone	Difference	95% Confidence Interval	Difference	95% Confidence Interval
Percent of households which were "identical repeat"	86.5	86.6	88.4	-	-	-	-
Percent of interviews conducted							
in a personal visit	78.3	96.3[1/]	19.6	-	-	-	-
by telephone	21.7	3.7	80.4[2/]				
Noninterview Rate (Percent)							
Type A	4.2	4.1	3.9	0.1	-0.3 to 0.5	0.3	-0.1 to 0.7
Type Z	1.9	2.0	2.4	-0.1	-0.3 to 0.1	-0.5*	-0.7 to -0.3
VICTIMIZATION RATES FOR SELECTED CRIMES AGAINST PERSONS (PER 1000 PERSONS 12 OR MORE YEARS OLD)							
Total Persons 12+							
Total Crimes Against Persons	129.4	130.2	119.1	-0.8	-8.4 to 6.8	10.3*	3.0 to 17.6
Crimes of Violence	32.0	34.1	31.8	-2.1	-6.2 to 2.0	0.2	-3.8 to 4.1
Aggravated Assault	9.6	11.7	10.1	-2.2	-4.6 to 0.2	-0.6	-2.8 to 1.7
Committed by stranger	6.2	6.3	4.9	-0.2	-1.9 to 1.6	1.3	-0.3 to 2.9
Committed by nonstranger	3.4	5.4	5.2	-2.0*	-3.6 to -0.4	-1.8*	-3.4 to -0.3
Crimes of Theft without Contact	94.5	92.8	84.6	1.7	-4.8 to 8.2	9.9*	3.6 to 16.2
Whites							
Total Crimes Against Persons	129.4	129.3	119.4	0.1	-7.9 to 8.2	10.0*	2.2 to 17.8
Crimes of Violence	30.7	31.8	31.6	-1.1	-5.3 to 3.1	-0.9	-5.1 to 3.3
Aggravated Assault	9.0	10.1	9.7	-1.2	-3.6 to 1.2	-0.7	-3.0 to 1.6
Committed by stranger	6.2	6.4	5.0	-0.2	-2.1 to 1.7	1.2	-0.5 to 2.9
Committed by nonstranger	2.8	3.7	4.6	-1.0	-2.4 to 0.5	-1.9*	-3.5 to -0.3
Crimes of Theft without Contact	96.3	94.4	85.5	1.9	-5.1 to 9.0	10.8*	4.0 to 17.5
Blacks							
Total Crimes Against Persons	130.5	146.0	116.9	-15.5	-39.5 to 8.5	13.6	-8.5 to 35.7
Crimes of Violence	42.4	56.4	33.6	-14.0	-29.6 to 1.6	8.8	-3.7 to 21.3
Aggravated Assault	14.5	26.2	13.0	-11.7*	-22.4 to -1.0	1.5	-6.3 to 9.3
Committed by stranger	6.2	6.6	2.0	-0.4	-5.9 to 5.1	4.2*	0.9 to 7.6
Committed by nonstranger	8.3	19.6	11.0	-11.4*	-20.6 to -2.1	-2.8	-9.8 to 4.3
Crimes of Theft without Contact	82.0	84.7	75.9	-2.7	-21.7 to 16.2	6.0	-12.2 to 24.2

*Indicates statistical significance at the 5 percent level.
[1/] Designated experimental procedure for the personal visit group.
[2/] Designated experimental procedure for the telephone group.

TABLE 2. MEAN LEVELS OF REPORTING FOR MAXIMUM PERSONAL VISIT AND MAXIMUM TELEPHONE TREATMENTS AND TESTS OF SIGNIFICANCE FOR DIFFERENCES BETWEEN REPORTED MEANS AND FOR DIFFERENCE BETWEEN POPULATION VARIANCES

Characteristic	Mean Reported Level (Per Person) [1]			Standard Deviation (Not standard error of sample mean)		
	Personal Visit Treatment	Telephone Treatment	t-Statistic	Personal Visit Treatment	Telephone Treatment	F-Statistic [2]
Number of Visits to Doctors	0.85	0.83	1.03	1.64	1.72	1.10**
Doctors' Charges Covered by Medicare ($)	11.78	11.90	-0.19	65.06	71.77	1.22**
Doctors' Charges Not Covered by Medicare ($)	3.86	3.72	1.00	14.80	15.82	1.14**
Number of Doctors' Visits Covered by Medicare	0.69	0.68	0.93	1.54	1.62	1.11**
Hospital Stays--Number of Times Doctors Seen	0.45	0.44	0.35	3.13	3.05	1.05**
Hospital Stays--Number of Doctors Seen	0.20	0.20	0.22	1.78	1.63	1.19**
Hospital Stays--Number of Days in Hospital	0.79	0.79	0.01	6.33	6.25	1.03*
Hospital Charges Covered by Medicare ($)	10.83	10.30	0.61	96.27	90.20	1.14**
Hospital Charges Not Covered by Medicare ($)	2.39	2.24	0.83	20.12	18.85	1.14**
Number of Hospital Visits Covered by Medicare	0.44	0.43	0.43	3.06	2.96	1.07**
Nursing Home Visits--Number of Times Doctors Seen	0.19	0.18	0.66	1.99	1.87	1.13**
Nursing Home Visits--Number of Doctors Seen	0.09	0.08	0.90	0.73	0.71	1.07**
Number of Days in Nursing Home	2.11	2.11	0.02	9.46	9.29	1.04**
Nursing Home Visits--Charges Covered by Medicare ($)	1.21	1.12	0.82	12.45	9.71	1.65**
Nursing Home Visits--Charges Not Covered by Medicare ($)	0.28	0.26	0.93	2.70	2.16	1.56**
Number of Nursing Home Visits Covered by Medicare	0.18	0.17	0.77	1.97	1.82	1.17*
Medical Services--Number of Uses	0.11	0.11	0.66	0.58	0.54	1.13**
Medical Services--Charges Covered by Medicare ($)	0.86	0.85	0.09	11.58	11.49	1.02
Medical Services--Charges Not Covered by Medicare ($)	0.20	0.20	0.12	2.55	2.54	1.00
Number of Medical Services Covered by Medicare	0.04	0.04	0.19	0.44	0.41	1.16**
Number of Prescriptions	1.51	1.49	0.84	2.19	2.24	1.04**

* Indicates statistical significance at the 5 percent level.

** Indicates statistical significance at the 1 percent level.

[1] Means in the telephone treatment are based on approximately 36,000 interviews and means in the personal visit treatment are based on approximately 18,000 interviews, and refer to the full eight month period of the experiment.

[2] F is the ratio of the larger to the smaller variance.