

SOME METHODOLOGICAL RESULTS FROM FOUR STATEWIDE  
TELEPHONE SURVEYS USING RANDOM DIGIT DIALING

Jane Williams Bergsten, Senior Statistician  
Sampling Research and Design Center  
Research Triangle Institute

INTRODUCTION

This paper will present some methodological results from four statewide telephone surveys which used a cluster random-digit-dialing procedure. The data presented include, for each of four states, estimated proportions of eligible clusters and estimated proportions of eligible telephone numbers within eligible clusters. Response rates are presented by such interviewer characteristics as previous experience, presence of Southern accent, and rapidity of speech. In addition, a learning curve is presented, and a comparison of response rates is made to determine the effect of permitting substitution of another telephone number when a refusal was obtained.

DESCRIPTION OF THE FOUR TELEPHONE SURVEYS

In the fall of 1978 Research Triangle Institute conducted post-election surveys in four western and midwestern states to determine attitudes and voting behavior associated with taxation and government spending.

A telephone survey of 300 randomly selected adult citizens was desired in each state. The interviewing was to begin the second day after the election, and continue for approximately two weeks.

The sample was selected in three stages: cluster, telephone number, and respondent. Using a data tape containing all operational telephone exchanges in the nation<sup>1</sup>, a listing of all operational telephone exchanges was produced for each of the four states included in the study. These listings served as first-stage sampling frames from which a simple random sample of 500 10-digit telephone numbers was selected for each of the four states. These selected telephone numbers would determine the selection of clusters and were therefore labeled "key" telephone numbers. Key telephone numbers that were found to be non-working, nonresidential, or out-of-state were so classified and were simply dropped from the study. Each key telephone number that was a working residential telephone number associated with an eligible household located in the designated state was included in the survey. In addition, such an eligible key number determined the selection of a cluster, and triggered the selection of additional telephone numbers within the cluster. A cluster was defined as all telephone numbers with the same leading 8 digits, and thus consisted of 100 similar telephone numbers, the final two digits of which ranged from 00 to 99.

Considering cost and distribution characteristics, the optimum allocation for each state was calculated to be 75 clusters of telephone numbers, each cluster consisting of 4 sample residential household telephone numbers.<sup>2</sup> Within a selected

cluster, additional telephone numbers, labeled "nonkey" telephone numbers, were therefore selected until a total of four eligible telephone numbers was discovered. The sample telephone numbers included in the study would therefore consist of the first four telephone numbers that were selected for the cluster and which met the eligibility criteria (see Figure 1).

The third stage of selection consisted of randomly selecting one adult in the household for interviewing. This was accomplished by assigning face sheets to eligible households in a prescribed manner. The interviewer listed the adult household members on the face sheet and numbered them in a prescribed way. The respondent selection

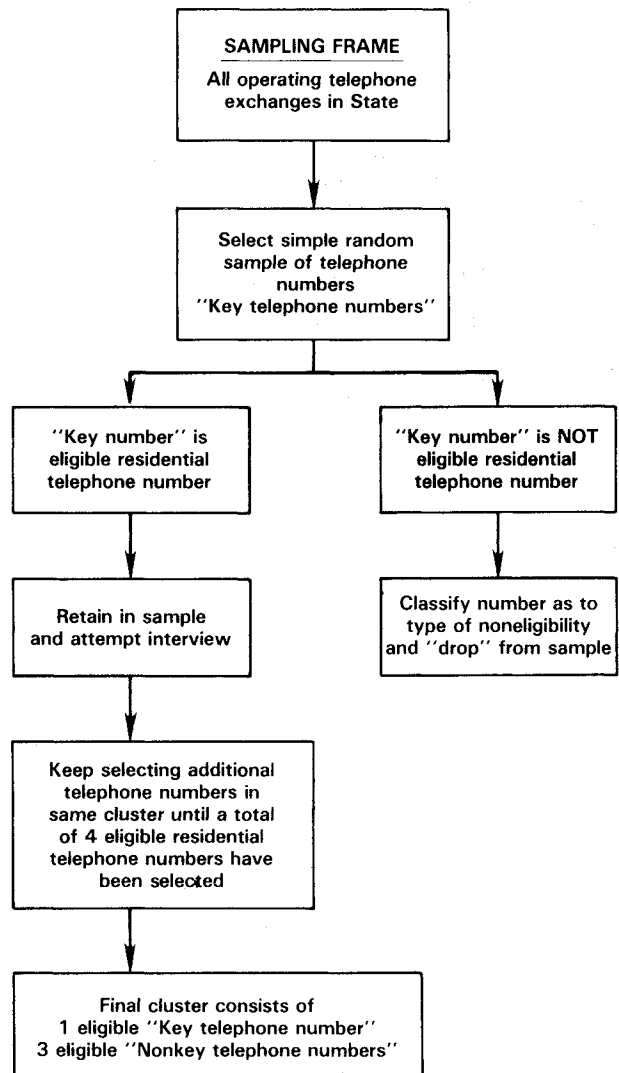


Figure 1. Selection Procedure.

table that was printed on the face sheet designated the randomly selected respondent. In all there were 12 different respondent selection tables.

A staff of 28 telephone interviewers was assigned to the study. The interviewers were telephone workers with considerable background in the areas of tracing previous respondents for follow-up surveys and of obtaining information about individuals by phone. Interviewing began immediately after the training session on November 9, 1978, and continued through November 27, 1978. Because of time zone differences, both night and day shifts were used so that interviewing could be conducted during a twelve-hour period 9 a.m.-9 p.m. in each of the four states. Interviewing was conducted on Saturdays, but not on Sundays and not during the Thanksgiving holiday.

A special "Result Code" was used by the interviewers to classify the result of each telephone call. The codes were grouped according to patterns of activity that would be associated with the result. For example, codes 21 or 22 referring to a busy number or a busy circuit result would permit the interviewer to make another attempt at that number in the same general time period. A 51 "no answer", on the other hand, had to be tried during a different time period, e.g., morning, afternoon, evening. The activity patterns were set up to insure against misclassification because of dialing errors or equipment failure and to insure wide time coverage in attempting to get a call through.

It had been our intention from the outset to adjust for noninterviews by selecting and interviewing additional telephone numbers in the same cluster. During the training period, however, the interviewers were told that additional attempts would be made to convert all refusals. About midway through the interviewing period when such additional attempts had proved to be fruitless, the substitution process was explained to the interviewers and put into operation. It was accomplished by selecting additional telephone numbers within the cluster, and then using the same respondent selection table as had been assigned to the noninterview. This latter step was taken in an attempt to keep to a minimum the biases associated with substitution for nonrespondents.

## RESULTS

### Eligibility Characteristics of Telephone Numbers

Table 1 shows the distribution of sample telephone numbers according to eligibility characteristics for each of the four states, for key telephone numbers, and for nonkey telephone numbers. The first-stage selections, consisting of 500 randomly selected telephone numbers per state, were activated in groups of 10 telephone numbers, in order of selection, until the desired number of clusters was obtained. This involved using slightly more than 300 selections in each of three states, namely Colorado, Michigan, and Oregon. In Idaho, all 500 first-stage selections were used, yielding only 58 clusters, rather than the desired 75. The reason for this can be seen in

**Table 1a. Disposition of Sample Key Telephone Numbers**

Eligibility Characteristics of Sample Telephone Number	Key Telephone Number			
	State			
	CO	ID	MI	OR
	Percent			
Residential	23.1	11.6	23.6	23.9
Nonresidential				
Not-in-service	66.2	75.6	55.2	68.1
Commercial	5.0	5.0	12.7	6.5
Subtotal	71.2	80.6	67.9	74.5
Out-of-State	0.3	0.4	0.3	—
Undetermined	5.3	7.4	8.2	1.6
Total	100.0	100.0	100.0	100.0
Number of telephone numbers	320	500	330	310
Number of clusters	74	58	78	74

**Table 1b. Disposition of Sample Nonkey Telephone Numbers**

Eligibility Characteristics of Sample Telephone Number	Nonkey Telephone Number			
	State			
	CO	ID	MI	OR
	Percent			
Residential	57.2	57.2	57.6	63.3
Nonresidential				
Not-in-service	22.9	25.3	25.3	22.2
Commercial	11.0	9.8	10.5	6.3
Subtotal	33.9	35.1	35.8	28.5
Out-of-State	—	—	0.1	—
Undetermined	8.9	7.7	6.5	8.2
Total	100.0	100.0	100.0	100.0
Number of telephone numbers	699	521	707	523
Number of clusters				

**Table 1c. Disposition of Sample Telephone Numbers by Type of Selection.**

Eligibility Characteristics of Sample Telephone No.	Type of Selection	
	Key Telephone No. Four States Combined	Nonkey Telephone No. Four States Combined
	Percent	
Residential	19.5	58.6
Nonresidential		
Not-in-service	67.3	24.0
Commercial	7.0	9.6
Subtotal	74.4	33.5
Out-of-State	0.3	0.0
Undetermined	5.9	7.8
Total	100.0	100.0
Number of telephone numbers	1,460	2,450
Number of clusters	284	

the first row of Table 1. Idaho had a much lower percentage of residential telephone numbers than did the other three states. While approximately 23 percent of the first-stage telephone numbers turned out to be working residential numbers in each of the other states, only half that percentage were so classified in Idaho. Note also that 76 percent of the Idaho first-stage numbers were "not-in-service" compared to 55 percent to 68 percent in the other three states. The larger percentage of commercial numbers in Michigan, 13 percent compared to rates half that size in the other three states, may be related to its heavy industrialization.

By comparing key telephone numbers with non-key numbers, one can see that the proportion residential among nonkey numbers is roughly three times that for key numbers. That is, about 59 percent of the nonkey telephone numbers were residential as compared to about 20 percent of the key numbers. It was, of course, this anticipated improved percentage, and the associated cost implications that had led us to use the cluster sample approach.

#### Telephone Call Characteristics

Telephone calls are distributed in Table 2 according to the outcome of the call, for key telephone numbers and for key and nonkey telephone numbers combined. Of all the telephone calls placed in the four-state survey, about one-fourth were answered telephone calls to eligible residences and another one-fourth yielded sufficient information to determine that the number should be dropped from the survey because of ineligibility. Almost half of the calls were not completed, primarily because of "no answer" (31 percent). When telephone calls made to key telephone numbers are examined, it can be seen that only about 13 percent were answered calls to residences. Another 43 percent of the calls yielded

Table 2. Result of Telephone Call by Type of Selection

Result of Telephone Call	Type of Selection	
	Key Telephone Number	All Telephone Numbers
	Percent	
Residence, call answered	13.0	25.3
Status determined, number dropped		
Number not-in-service	37.9	22.6
Commercial number	2.5	3.2
Incorrect number (bridging)	2.9	2.1
Subtotal	43.4	27.9
Calls to business office	1.3	1.4
Call not completed		
Poor connection	0.8	0.8
Busy (fast or slow)	13.6	9.5
Failed to go through	8.1	4.2
Temporarily out of order	0.3	0.1
No answer	19.6	30.8
Subtotal	42.4	45.3
Total	100.0	100.0
Total number of telephone calls	4,284	10,799

Table 3. Number of Calls Needed to Obtain Specified Status

Number of Calls	To Obtain an Interview	To Determine Eligibility	
		All Eligibles	All Noneligibles
Percent			
1	29	59	29
2	22	17	45
3	17	11	9
4	11	5	7
5	7	4	5
6	5	2	2
7	3	1	2
8+	6	1	1
Total	100	100	100
Total number of telephone calls	1,102	1,597	1,853
Mean number of calls	3.01	1.94	2.33

information sufficient to classify the associated number as dropped from the survey due to ineligibility and an additional 42 percent of the calls were not completed. The key numbers, therefore, can be characterized as having a higher proportion of not-in-service calls, a lower proportion of answered residential calls, and a lower proportion of no-answer calls.

Table 3 shows percentage distributions and means for the number of telephone calls necessary to obtain an interview or to determine the eligibility status of the telephone number. For all telephone numbers for which an interview was eventually obtained, the total number of calls averaged approximately 3. Among all telephone numbers that were eventually classified as "eligible working residential telephone numbers" the average number of calls needed, in order to establish the fact that the number was indeed eligible, was about 1.9 calls. For all telephone numbers that were eventually classified as ineligible, an average of 2.3 calls was required in order to establish the ineligibility.

#### Survey Response Rates

The response disposition for the survey as a whole is presented in Table 4. The response categories were set up to be as comparable as possible to those used in other papers. The overall re-

Table 4. Response/Nonresponse Components for Total Telephone Sample.

Disposition	n	Percentages Including Undetermined	Percentages Excluding Undetermined
		Percent	
Completed interviews	1,102	56.0	65.1
Partial interviews	20	1.0	1.2
Refusal by R	134	6.8	7.9
Other refusal	370	18.8	21.9
Other noninterview	20	1.0	1.2
R absent	47	2.4	2.8
Undetermined status	277	14.1	—
Total	1,970	100.0	100.0

sponse rate was 56 percent, if one considers all of the numbers with undetermined status to be eligible residential numbers. If these latter numbers are eliminated, and classified as ineligible, the response rate was about 65 percent. Note, that approximately one-fourth of the numbers resulted in a refusal, and the vast majority of these refusals were at the household level, rather than by the selected respondent.

The low response rate was, of course, the cause of great concern. There was a variety of possible reasons for it, among which are the following:

- 1) Relative inexperience of interviewers.
- 2) Need for obtaining a household roster.
- 3) Excessive length of interview, which averaged somewhat over 22 minutes.
- 4) Technicality of subject matter and complexity of questions.
- 5) Extremely short interviewing period.

A slightly different method of computing response rates was used in Table 5, in addition to the two methods used in Table 4. Row 8 of Table 5 shows the response rate calculated by dividing the number of completed interviews by the number of completed interviews plus the number of refusals. Included as refusals are both re-

spondent and household refusals, and also refusals obtained after completion of only part of the interview. This type of response rate was felt

**Table 5. Response Rates by Type of Selection**

Disposition	Type of Selection			Total Sample
	Key Telephone Number	Nonkey Telephone No.		
	Original Selection	Substitute Selection		
Frequency				
(1) Interview	163	564	375	1,102
(2) Refusal, any type	90	257	176	523
(3) Interviews + refusals = (1)+(2)	253	821	551	1,625
(4) Non-int. or partial int., other	23	18	27	68
(5) Total known eligibles = (3)+(4)	276	839	578	1,693
(6) Undetermined eligibility	88	103	86	277
(7) Total = (5)+(6)	364	942	664	1,970
Response rate Percent				
(8) (1)÷(3)	64.4	68.7	68.1	67.8
(9) (1)÷(5)	59.1	67.2	64.9	65.1
(10) (1)÷(7)	44.8	59.9	56.5	55.9

**Table 6. Response Rates by Selected Interviewer Characteristics\***

Selected Interviewer Characteristics	# Ints.	# Refs.	Total	Percent Resp. Rate	Sig. Level of Diff. $\alpha^{**}$
Previous Interviewing Experience					
1. Less than 6 months or none	710	344	1,054	67.4	.032
2. 6 months or more	368	110	478	77.0	
Total	1,078	454	1,532	70.4	
Accent					
1. Strong southern	308	179	487	63.2	
2. Mild southern	254	92	346	73.4	
3. Midwestern neutral	429	147	576	74.5	
4. Mild northeastern	71	23	94	68.3	
5. Strong northeastern	16	13	29	55.2	
Total	1,078	454	1,532	70.4	
1,2. Southern	562	271	833	67.5	.075
3,4,5. Other	516	183	699	73.8	
Total	1,078	454	1,532	70.4	
1,5. Strong	324	192	516	62.8	.005
2,3,4. Mild or neutral	754	262	1,016	74.2	
Total	1,078	454	1,532	70.4	
Rapidity of Speech					
1. Rapid	442	176	618	71.5	
2. Medium	343	152	495	69.3	
3. Slow	293	126	419	69.9	
Total	1,078	454	1,532	70.4	

\* There were 28 telephone interviewers. Interviewer identification was missing from 24 interviews and 69 refusals. These 93 cases were excluded from this table.

\*\*Using a t-test and assuming an unrestricted random sample of 28 unequal size clusters.

to provide a more sensitive measure of interviewer effectiveness than the other methods of calculation, and was used in the remaining tables and graphs.

As can be seen in Table 5, this response rate was calculated to be 64.4 percent for key telephone numbers. This was slightly lower than the response rates for the other two groups of nonkey numbers, and probably reflects the lower effectiveness of the interviewers at the early stages of the data collection period. Note that the response rates for the original-selection nonkey numbers and for the substitute nonkey numbers are each approximately 68 percent. We had feared that permitting the interviewers to select substitute telephone numbers as replacements for nonresponses would result in an increased refusal rate. Our thinking was that there would be less incentive to try to convince the uncooperative to participate in the survey, when another number might prove to be more cooperative. Our fears seem to have been unfounded.

#### Response Rates by Interviewer Characteristics

More experienced interviewers achieved a higher overall response rate than less experienced interviewers. Those with 6 months or more previous interviewing experience obtained an average response rate of 77 percent compared to a rate of 67 percent for those with less than 6 months or no previous experience (Table 6). This difference could be explained because of a positive association between increased experience and increased effectiveness or because of higher maintenance rates among more effective interviewers.

The interviewer's accent was also related to the achieved response rate. Because of RTI's southern location as contrasted with the western and midwestern location of the four survey states, we were particularly interested in determining whether presence of a southern accent had any affect on the response rate. It did indeed appear to have had a slight effect. Interviewers with southern accents had an average response rate of approximately 68 percent compared to an average response rate of 74 percent for interviewers not having a southern accent. A difference of 11 percentage points can be noted when the response rates are compared for interviewers with a strong accent and those with only a mild accent or none at all. This latter group achieved a response rate of 74 percent compared to a rate of 63 percent for those with a strong accent.

The interviewers were also classified according to how rapidly they spoke. The response rates were consistently about 70 percent for each of the rapidity-of-speech categories.

#### Effect of Learning on Response Rate

The response rate is shown by survey day in Figure 2. It increased rather steadily for the first week of interviewing, from about 58 percent on the first day to 76 percent on day 6. It then dropped slightly and remained at about 70 percent

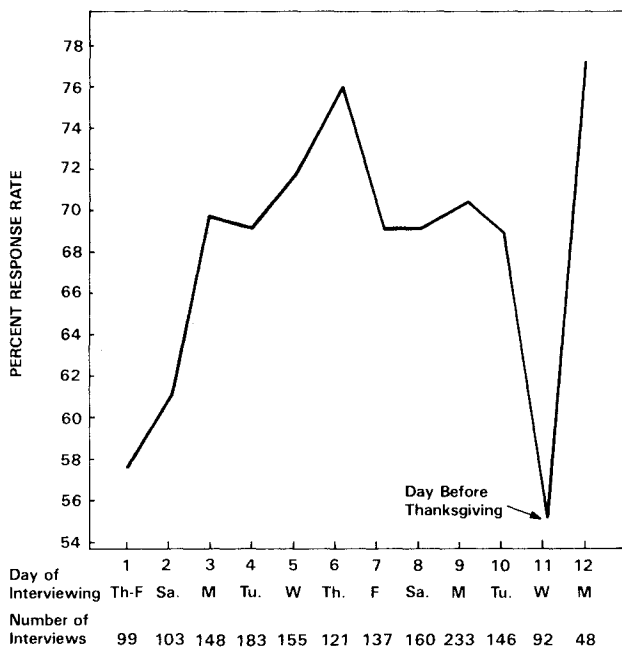


Figure 2. Response Rates by Day of Interviewing.

for four days. It dropped severely to a low of 55 percent the day before Thanksgiving, indicating perhaps a greater focus on preparation for Thanksgiving dinner than on taxation and spending considerations. The response rate peaked again on the last day of interviewing reaching a high of 77 percent.

#### SUMMARY AND CONCLUSIONS

There is evidence that the proportion of residential telephone numbers can vary substantially from state to state, and that cluster sampling can indeed substantially increase the proportion of telephone numbers that yield eligible residential households.

Our four-state experience indicated that eligibility status can be determined by an average of about two telephone calls, and an interview can be obtained in an average of one additional telephone call.

While both interviewer experience and absence of a strong accent appear to be associated with a higher response rate, there is no indication that rapidity of speech and response rate are associated.

#### FOOTNOTES

1. This tape is available from the American Telephone and Telegraph Company's Long Lines Department.
2. See Waksberg, J. Sampling Methods for Random Digit Dialing, Journal of the American Statistical Association, March 1978, pp. 40-46.