SOME COMPARATIVE EFFECTS OF PERSONALIZED AND NONPERSONALIZED PROCEDURES IN SURVEYS MAILED TO TEACHERS

Glen Robinson, National Education Association
Simeon P. Taylor, III, National Education Association

Introduction

Since 1959, the Research Division of the National Education Association has been engaged in a concerted effort to streamline its statistical and data-gathering procedures. This effort, termed the Sampling Project, has brought together the most up-to-date sampling procedures available and is adapting them to teacher, principal, and superintendent populations. The central purpose of the project is to improve the accuracy and speed of data collecting and reporting in nationwide surveys.

A well-known and important factor affecting the accuracy of data collected through mailed surveys is the usually low rate of response and the possible accompanying high nonresponse bias. The initial phase of the Sampling Project focused on this problem. Through careful planning, efficient control of mailing lists, appropriate processing of incoming questionnaires, effective initial contact letters, and effective follow-up procedures, remarkable response rates have been achieved. The mean net response rate (useable questionnaires returned) for seven surveys conducted in 1960 and 1961 was 95.7 percent.1/

In the first phase of the project, the objective was to secure the highest response rates that could possibly be obtained from the populations sampled; therefore, cost was not a factor considered in initial contact and follow-up procedures. The many theories on methods of increasing response rates were examined and the most plausible were incorporated into the survey procedures.

Since it was believed that personalized approaches in initial contact letters and follow-up procedures were essential to achieving high response rates, such approaches were employed to the fullest. For example, survey instruments were accompanied by a personally addressed and typed letter, signed by the Director of the Research Division. This letter explained the nature of the survey and the importance of each teacher's response. This letter contained a copy of the survey instrument and a self-addressed envelope bearing a postage stamp.

The general pattern of follow-up procedures consisted of a personal letter sent via air mail to nonrespondents at the end of two weeks. A telegram follow-up was sent at the end of four weeks, and another at the end of five weeks. Cut-off dates for receipt of retured questionnaires were established at eight weeks following the initial mailing date.

In conducting the surveys, special effort was made to glean as much information as possible about the populations studied and their response patterns. This information has been valuable in analyzing factors associated with nonresponse and the accompanying biases.

The Cost-Quality Problem

The estimate of $214 per 100 persons for typing, postage, and telegraphy used in the highly personalized contact and follow-up procedures of this first phase of the project made the procedures prohibitive for general use in nationwide surveys. The Division, therefore, was faced with the problem of how much the highly personalized approach could be relaxed without causing a significant drop in the response rate. We believed the rate should be kept above 90 percent.

Several questions had to be answered. Among them were:

1. How much of the high response rate, if any, was due to the personalized effect of the telegram?
2. Could the telegrams be replaced by letters and still achieve a response rate above 90 percent within a reasonable time?
3. Were personalized initial contact and follow-up letters essential for a minimum response rate or could nonpersonalized letters be substituted?

To find answers to these questions, an experimental design was developed to test variations in initial contact and follow-up procedures.

Relative Drawing Power

Two hypotheses were formulated for the purpose of testing the relative drawing power of personalized and nonpersonalized contact procedures. These hypotheses were:

1/ For a description of the seven surveys, the procedures used, and estimates of the sampling variability in the studies see: Robinson, Glen, and McCall, Chester H., Jr. "Some Statistical Findings from Nationwide Teacher Polling." Proceedings of the Social Statistics Section, 1961. Washington, D. C.: American Statistical Association, 1961. p. 56-63.

I. Letter follow-up procedures have the power to elicit from samples of teacher populations in mailed surveys a rate of response equal to that elicited by procedures involving telegrams.

II. Personalized initial contact letter and follow-up procedures have the power to elicit a significantly higher rate of response from teacher populations in mailed surveys than do nonpersonalized letters.

The Experimental Design

The NEA Teacher Opinion Polls were selected as the test media. The survey instruments used in these polls consist of a single sheet of paper printed front and back containing approximately 14 opinion questions and 12 status questions. These instruments are mailed to a nationwide probability sample of classroom teachers. Samples have ranged in size from 1,147 to 1,633.

Polls I and II were conducted in the spring of 1960 with almost identical contact and follow-up procedures. These consisted of an initial contact letter and a first follow-up letter personally addressed, typed, signed, and sent air mail, and second and third follow-ups with telegrams. The polls achieved responses of 96.8 percent and 94.9 percent, respectively, within an eight-week period.

Poll III was conducted in the spring of 1961. The follow-up procedure was altered by the substitution of a second personalized letter for the first telegram follow-up. Figure I shows the variations in the response patterns among the three polls. Table 1 shows the cumulative response rates. It should be noted that Poll III achieved a 96.6-percent response within eight weeks. Obviously, the variations among the response rates of Polls I, II, and III could be attributed to chance alone.

When Poll IV was conducted in the spring of 1962, it was decided to depart substantially from the procedures in the previous polls as follows:

1. The nationwide probability sample of 1,464 classroom teachers selected for Poll IV was randomly divided into two groups of 732.[2/]

2. Neither group received telegrams. Both groups received an initial contact letter followed by a postal card reminder at the end of two weeks and four follow-up letters at two-week intervals. The contents of the letters and the postal cards were identical for the two groups. Follow-ups for the two groups were mailed simultaneously.

3. The treatment given the two groups differed in that the contact and follow-up letters sent to Group A were personally addressed to the teacher, typed, and signed by the Director of the Division; whereas the contact letters and the first two follow-up letters sent to Group B were impersonally addressed, duplicated by offset machine, and bore only a facsimile signature. The last two letters sent to Group B were personalized for the reason explained later.

The two purposes for this division were:

1. To compare the results received from Group A with Polls I, II, and III to determine

2/ For description of sampling procedure see footnote 1.

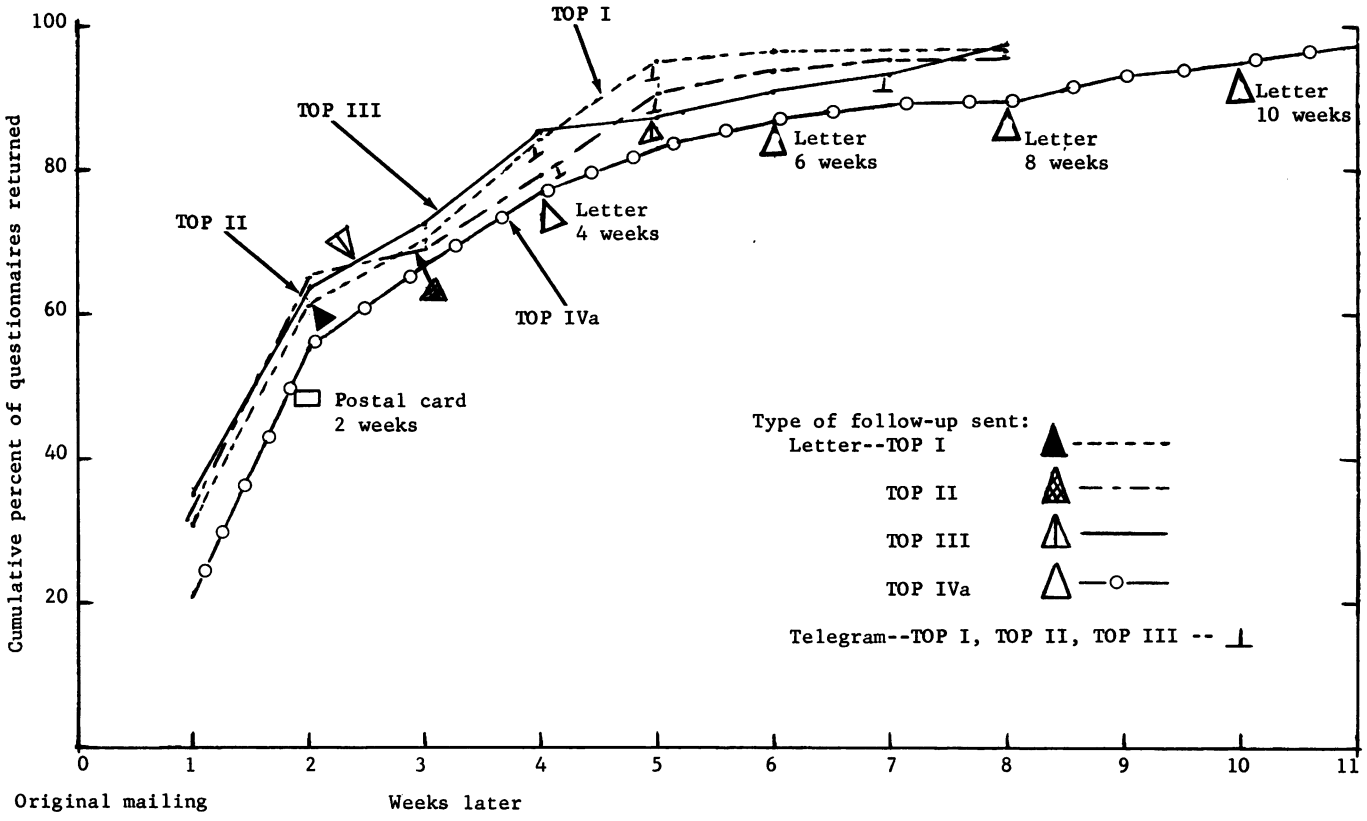
TABLE 1.--CUMULATIVE PERCENT OF RESPONSE[a/] FOR TEACHER OPINION POLLS I, II (1960), AND III (1961), AND FOR TEACHER OPINION POLL IV, EXPERIMENTAL GROUP A (1962)

Weeks after mailing	TOP I n=1,149	TOP II n=1,147	TOP III n=1,633	TOP IV[a/b/] n=732
1	2	3	4	5
1	30.1%	31.5%	35.1%	20.9%
2	61.4	64.8	64.0	55.3
3	71.1	69.4	72.2	66.8
4	84.1	79.7	84.2	76.5
5	95.3	91.2	86.8	82.9
6	96.3	94.0	90.9	86.3
7	96.8	94.9	93.3	88.7
8	96.8	94.9	96.6	89.6
9	...	...	...	93.0
10	...	...	...	94.4
11	...	...	...	96.4

a/ Percents of response are cumulative net responses which include only those persons returning useable questionnaires.

b/ These data are for TOP IV experimental Group A only.

FIGURE I--RESPONSE PATTERNS FOR TEACHER OPINION POLLS I (1960), II (1960), III (1961), AND IVa (1962)



if telegrams could be eliminated completely, and thus serve as a test for Hypothesis I.

2. To have Group B serve as a control for Group A in testing Hypothesis II (personalized vs. nonpersonalized letters).

Experimental Findings

A brief discussion of the findings of these experiments follows:

Telegrams vs. Personalized Letters

Figure I shows the response pattern for Poll IV, and Group A. Table 1 shows the cumulative net response ratios. Comparison of response curves in Figure I indicate how the response to the personalized letter by Group A differs from that obtained in the three previous polls. Utilization of the Kolmogorov-Smirnov test demonstrates that the response distribution for Poll IV, Group A, was significantly different from the other three polls (α=.001). At the end of eight weeks the response rate for Group A was only 89.6 percent compared with a mean response of 96.1 percent for the other three polls. This difference was significant at the .05 level.

But the important thing is that at the end of 11 weeks the response rate was 96.4 percent. This was 1.5 percentage points above the 94.9-percent response in Poll II and only 0.4 percentage points below Poll I, both of which had used two telegram follow-ups. Obviously, variations of this size could easily occur by chance alone.

TABLE 2.--TREATMENT OF GROUPS A AND B IN TEACHER OPINION POLL IV

Action and date	Group A: Experimental, personalized treatment (n = 732)	Group B: Control, nonpersonal treatment (n = 732)
1	2	3
Initial contact February 7	Personalized letter	Nonpersonalized letter
1st follow-up Two weeks	Nonpersonalized postal card reminder	Nonpersonalized postal card reminder
2nd follow-up Four weeks	Personalized letter	Nonpersonalized letter
3rd follow-up Six weeks	Personalized letter	Nonpersonalized letter
4th follow-up Eight weeks	Personalized letter	Personalized letter
5th follow-up Ten weeks	Personalized letter	Personalized letter

TABLE 3.--VARIATIONS IN FIRST EIGHT WEEKS OF FOLLOW-UP

Factor varied	Group A (Experimental)	Group B (Control)
1	2	3
Method of writing	Wording of entire letter written on automatic electric typewriter	Wording of entire letter typed on multilith mats and duplicated on an offset machine
Inside address	Addressed by name and school address to the individual in the sample	No inside address
Salutation	Salutation contained the name of the individual recipient, i.e., Dear Mr. Doe.	All letters bore general salutation of "Dear teacher"
Signature	Personal signature of the director of the NEA Research Division in ink	Facsimile signature of director of the NEA Research Division from multilith mat

Since the final response rate for Poll IV, Group A, was substantially equal to the rates obtained in the previous polls, and since 11 weeks is a reasonable period of time in these studies, Hypothesis I is acceptable. Hence, we conclude that personalized letter follow-up procedures have the power to elicit from teacher populations, within a reasonable time, a response equal to that elicited by procedures involving telegrams.

There is evidence, however, that the use of telegrams in follow-up procedures does reduce the time of response.

Personalized vs. Nonpersonalized Letters

Hypothesis II was tested by a carefully controlled experiment, the general design of which was briefly described previously. A probability sample of 1,464 classroom teachers, drawn from the nearly 1.5 million teachers in the nation, was used for Teacher Opinion Poll IV conducted in the spring of 1962.

This sample was randomly divided into two groups. For an eight-week period, Group A received personalized treatment and Group B received nonpersonalized treatment. At the time the experiment was designed, we believed that the power of nonpersonalized letters to elicit responses was substantially less than that of personalized letters. Because of this belief, a safety factor was added to Group B to assure a high terminal response rate. At the end of eight weeks and at the end of ten weeks nonrespondents in Group B as well as Group A were to receive personalized letters.

Although the personalized treatment for Group B after eight weeks was included as a safety factor to assure useable opinion data, it was believed that any substantial closing of the response gap that existed between the two groups at the end of eight weeks and at the end of the experimental period would be additional evidence on the relative power of personal and impersonal treatments.

The treatment applied to the two groups are given in Table 2.

The contents of all communications was the same for both groups. The outside address of the envelope contained the name and school address of the individual teacher in both groups. All mailings were the same for both groups. All postage stamps were identical for both groups. The four factors which were varied during the first eight weeks are shown in Table 3.

Results of Personalized vs. Nonpersonalized

The experimental Hypothesis II was that personalized initial contact letters and follow-up procedures have the power to elicit a significantly higher response rate from teacher populations in mailed surveys than do nonpersonalized letters.

The statistical or null hypothesis was that there was no difference between the response pattern shown by Group A and the response pattern shown by Group B.

Figure II shows graphically the response for the two groups.

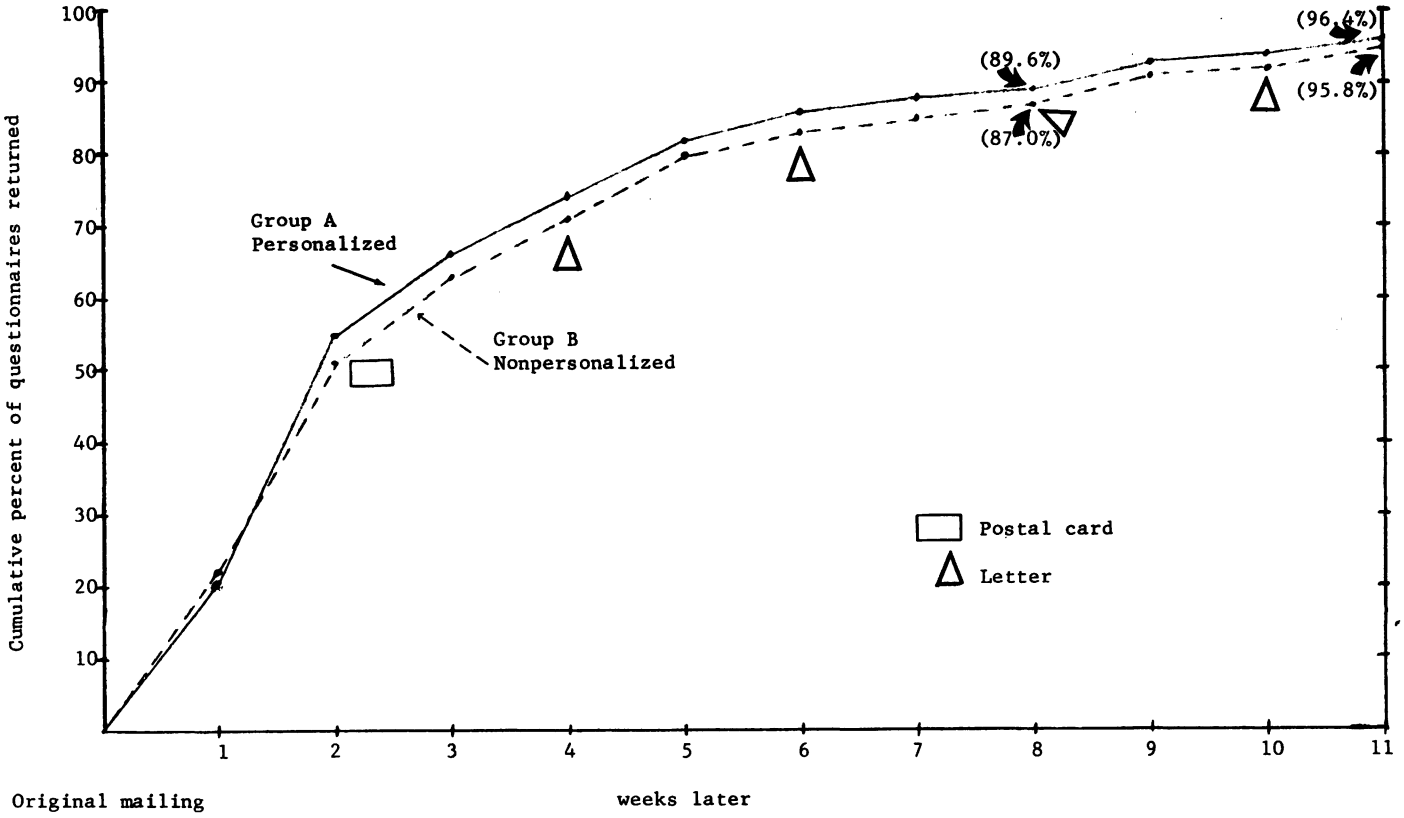
Table 4 presents cumulative weekly responses by number and percent for the two groups during the 11-week period.

A comparison of the response data for the experimental and control groups reveals the following:

TABLE 4.--CUMULATIVE RECEIPTS AND PERCENT OF RESPONSE FROM INITIAL MAILING AND SUCCESSIVE FOLLOW-UPS BY WEEKS FOR TEACHERS RECEIVING PERSONALIZED AND NONPERSONALIZED LETTERS, TEACHER OPINION POLL IV, GROUPS A AND B

Weeks from initial mailing	Group A (Experimental) n = 732		Group B (Control) n = 732		Percentage point difference in predicted direction (A > B)
	Number	Percent	Number	Percent	
1	2	3	4	5	6
1	153	20.9%	163	22.3%	-1.4
2	405	55.3	379	51.8	3.5
3	489	66.8	467	63.8	3.0
4	560	76.5	539	73.6	2.9
5	607	82.9	586	80.1	2.8
6	632	86.3	614	83.9	2.4
7	649	88.7	626	85.5	3.2
8	656	89.6	637	87.0	2.6
9	681	93.0	670	91.5	1.5
10	691	94.4	677	92.5	1.9
11	706	96.4	701	95.8	0.6

FIGURE II--CUMULATIVE RESPONSES FROM TEACHERS RECEIVING PERSONALIZED AND NONPERSONALIZED INITIAL AND FOLLOW-UP LETTERS IN TEACHER OPINION POLL IV, GROUPS A AND B



1. With the exception of the first week, the response rate for the experimental group consistently exceeded that of the control group. This factor would suggest that the personalized treatment was affecting response in the predicted direction.

2. However, these differences were surprisingly small, ranging from 2.4 to 3.5 percentage points in the predicted direction during the initial 8-week period. The average percentage point difference was 2.9 for the eight weeks. When these differences were subjected to the Kolomogorov-Smirnov test, it was shown that differences this great could be expected to occur by chance alone at well above the .10 level. Thus, there is not sufficient evidence to cause us to reject the null hypothesis.

3. Therefore, we must reject, with reservations, experimental Hypothesis II and conclude that in studies of this type personalized letters do not significantly affect the response rates exhibited by samples of teacher populations. In reaching this conclusion, there are several conditioning factors which will be discussed later.

4. At the end of the eighth week, the difference between the experimental and the control groups was 2.6 percentage points in the predicted direction. At this point the control group began to receive the same personal treatment as the experimental group. Within the three weeks that followed, the percentage point gap between the control and the experimental groups was closed to the point that the control group lagged by only 0.6 of a percentage point. Although these observations are interesting and are consistent with the slight difference shown earlier, the difference is so small that it is attributable to chance.

Some General Conclusions and Observations

From our experimental work in the field of sample surveys mailed to teacher populations, we draw the following general conclusions and make these observations:

1. One of the criticisms frequently directed at those who use mailed survey techniques is the generally poor response. The NEA Research Division's Sampling Project has overcome this criticism through its demonstration that consistently high response rates can be obtained from probability samples of teacher populations.

2. Our experience demonstrates that high response rates (above 90%) can be obtained at a reasonable cost. We estimate the costs for mailing and typing the personalized initial and follow-up letters described for Group A to be $30 per 100 persons in the sample. We estimate the cost of mailing and duplicating of the nonpersonalized letters and the two personalized letters described for Group B to be $17 per 100 persons in the sample.

3. We find no evidence that the use of telegrams as a medium for follow-up communications with teacher samples increases the final response rates where time is not a factor.

4. We find evidence that the use of telegrams tends to hasten the response of teacher samples.

5. In this study the differences between the response pattern of the samples of teachers receiving personalized and nonpersonalized letters was not statistically significant.

Similar findings were reported by Clausen and Ford[3/] in studies of the responses of World War II veterans to mailed surveys on attitudes toward, and information about, National Service Life Insurance. They reported:

> In mail follow-ups of veterans who had not responded to the initial questionnaire, personalized salutation and true signature did not lead to significant increases over nonpersonalized forms in rate of response....

Although the findings of our study failed to show a significant difference in the response patterns as between personalized and nonpersonalized approaches, we would caution against generalizing to surveys that differ either in type of population studied or in nature of survey instrument used.

The survey instruments used in these studies were short, multiphasic questionnaires. Their content covered several subjects of great interest to teachers generally. Although there is evidence from our experience and from the experience of other researchers that the length of the questionnaires does not have a significant effect on response rates, there is evidence that multiphasic questionnaires tend to have a greater drawing power for response than do monophasic questionnaires.

Clausen and Ford[4/] found:

> A multiphasic survey, covering several potentially interesting topics, yielded higher rates of response than a single subject survey of the same population, and also greatly lessened an interest bias in response.

Quite possibly the reason for the low differential in drawing power between personalized and nonpersonalized letters found in our study is an overriding high-interest effect. It could be reasoned that in a study in which the survey instrument is monophasic and confined to a single area of inquiry of low interest to teachers, the superior drawing power of personalized letters might be significantly demonstrated. In studies presently being planned, we hope to test this possibility.

3/ Clausen, John A., and Ford, Robert N. "Controlling Bias in Mail Questionnaires." Journal of the American Statistical Association 42: 497-511; December 1947.

4/ Ibid., p. 497.

6. It is our tentative conclusion from our experience to date that an optimum contact and follow-up procedure where cost-quality relationships are concerned and where time is not a crucial factor, consists of a combination of the personalized and nonpersonalized approaches. Such a procedure might consist of a personalized initial letter, nonpersonalized postal card reminder, nonpersonalized letter, and a personalized letter sequence. In our opinion the optimum time between contacts seems to be one to two weeks. We are planning studies to test the validity of these assumptions.