R E P O R T R E S U M E S

ED 015 887 SP 001 377

THE SELECTION OF A NATIONAL RANDOM SAMPLE OF TEACHERS FOR EXPERIMENTAL CURRICULUM EVALUATION.

BY- WELCH, WAYNE W. AND OTHERS

PUB DATE 10 NOV 67

EDRS PRICE MF-$0.25 HC-$0.64 14P.

DESCRIPTORS- CONTROL GROUPS, *EVALUATION TECHNIQUES, EXPERIMENTAL GROUPS, NATIONAL SURVEYS, *PHYSICS CURRICULUM, *PHYSICS TEACHERS, PROGRAM COSTS, QUESTIONNAIRES, RESEARCH DESIGN, *RESEARCH METHODOLOGY, RESEARCH PROBLEMS, *SECONDARY SCHOOL TEACHERS, SELECTION, STATISTICAL DATA, TABLES (DATA), HARVARD PROJECT PHYSICS, SCIENCE TEACHERS ASSOCIATION

MEMBERS OF THE EVALUATION SECTION OF HARVARD PROJECT PHYSICS, DESCRIBING WHAT IS SAID TO BE THE FIRST ATTEMPT TO SELECT A NATIONAL RANDOM SAMPLE OF (HIGH SCHOOL PHYSICS) TEACHERS, LIST THE STEPS AS (1) PURCHASE OF A LIST OF PHYSICS TEACHERS FROM THE NATIONAL SCIENCE TEACHERS ASSOCIATION (MOST COMPLETE AVAILABLE), (2) SELECTION OF 136 NAMES BY A TABLE OF RANDOM NUMBERS DERIVED FROM ORDINAL NUMBERING OF THE 16,702 LISTINGS FOR THE CONTINENTAL U.S., (3) SENDING OF A DESCRIPTIVE INVITATION TO PARTICIPATE IN EXPERIMENTAL EVALUATION OF THE COURSE (INCLUDING A QUESTIONNAIRE TO INDICATE WILLINGNESS, REASON FOR REFUSAL, AND INFORMATION ON TEACHING LOAD, TEXT USED, ETC.), (4) COMPARISON OF ACCEPTORS WITH NON-ACCEPTORS (ACCEPTORS APPEARED MORE RECEPTIVE TO INNOVATION), (5) ASSIGNMENT BY RANDOM NUMBERS OF 46 TEACHERS TO THE EXPERIMENTAL GROUP AND 26 TO THE CONTROL GROUP (ATTRITION OVER THE FOLLOWING 4 MONTHS REDUCED THE EXPERIMENTAL GROUP TO 36 AND THE CONTROL GROUP TO 21). POSSIBLE BIAS IS DISCUSSED, LEADING TO CHARACTERIZATION OF THE POPULATION AS "ALL HIGH SCHOOL PHYSICS TEACHERS ON THE 1966 NSTA LIST WHO DID NOT MAKE SUMMER COMMITMENTS MORE THAN 3 MONTHS IN ADVANCE AND WHO WOULD NOT BACK OUT OF THE RESEARCH AGREEMENT." COST AND COMPARISON WITH A VOLUNTEER GROUP ARE ALSO DISCUSSED. (AF)

The Selection of a National Random Sample of Teachers for Experimental Curriculum Evaluation

Wayne W. Welch, Herbert J. Walberg,
and Andrew Ahlgren
Harvard University

Introduction

Given the task of evaluating the effects of a new high school physics course, the evaluation group of Harvard Project Physics* attempted to (a) select the participating physics teachers from the national population on a truly random basis and (b) to assign them randomly to experimental and control groups. The random selection was particularly appropriate for the Project Physics course because it is intended for a very diversified audience. Randomness justifies statistical inferences to a much larger population than is reasonable for a group of volunteer teachers.

In this paper we will describe what is to our knowledge the first attempt to select a national random sample of teachers. The discussion of the problems encountered and the success and possible value of the undertaking may be helpful to other curriculum research and evaluation groups.

*For a description of the materials and philosophy of Harvard Project Physics, see The Physics Teacher, 5, 5 (May 1967), pp. 197-233.

A list of the names and addresses of 16,911 physics teachers was purchased at a cost of $255 from the National Science Teachers Association (NSTA), a group which maintains the U. S. Registry of Junior and Senior High School Science and Mathematics Teaching Personnel. The NSTA has reported (1965) that the list is compiled from responses received from 81 percent of all secondary schools in the United States. We were unable to find a more complete, or even comparable, list from any other source.

Because of travel costs for teacher training, we limited our population to the 16,702 physics teachers listed for the continental United States. Numbers were assigned to each of the teachers according to their ordinal position on the list and a table of random numbers was used to select a total of 136 names. These two steps took a clerical assistant approximately four days to complete. (Actually the selections were made in two blocks. Initially 85 names were selected and the teachers contacted. Then to complete the roster of teachers available for summer training, a second random selection of 51 names was made.)

A registered letter describing the curriculum project was sent to the teachers together with an invitation to participate in an experimental evaluation of the course.

They were informed that those teachers agreeing to participate would be randomly assigned to one of two groups--an experimental group and a control group. Thus the teacher would have to be willing to be in either group if he consented to participate at all.

A description of the responsibilities of both groups was given in the letter. In summary, the experimental group would attend a six-week training session, take a series of tests, teach the course during the academic year 1967-68, and administer pre, mid, and post tests to their physics students. The control group would attend a two-day briefing session, take a series of tests, and administer the same battery of tests to their students; but they would continue to teach their regular physics courses. Travel expenses, summer school stipends, and course materials were to be provided by the curriculum project.

For reasons of expense and on the basis of a research design that was tried out with volunteer teachers during 1966-67, we decided to assign 40-45 teachers to the experimental group and 20-25 teachers to the control group. Thus it was planned that the evaluation would include 60-70 teachers with an estimated 3,000 to 4,000 students.

The teachers contacted were requested to return a one-page questionnaire, indicating either their willingness to participate in the experiment or their reasons for refusal. To take advantage of the opportunity to query a random sample of physics teachers, the questionnaire included several questions about school size, teaching load, physics enrollment, and text used. Several of these findings are reported by Welch (1967).

The invitation letters were sent by registered mail and replies were requested within two weeks. Uncertain as to what response rate to expect, we mailed 85 invitations. Responses were received from 68 within two weeks. Six registered letters were returned unopened with various post office indications: "addressee unknown," "unclaimed," or "moved--left no address." The remaining 17 were telephoned. Of these, six said yes, nine said no, and we were unable to locate two because of incorrect addresses. Thus in our first mailing, 77 teachers were contacted; 46 agreed to participate; 30 had prior commitments which prevented their acceptance; and 3 were not interested. Because the response rate of the first mailing was about 60%, a second mailing of 51 invitations was made. We hopedthis would give us approximately 70 teachers for the study.

A similar pattern of response, telephone calls, acceptance and rejection followed the second mailing. It is convenient here to talk about the group as a whole, although nearly five weeks elapsed between the first mailing (January 3) and the last teacher reply (February 9).

A total of 136 letters of invitation were mailed, but only 124 teachers actually were contacted. Nine letters were "returned to writer" by the post office and three others could not be reached by phone. These twelve names were omitted from the sample. Replies were received from 87 of the 124 teachers within two weeks after mailing. The remaining 37 were contacted by telephone to determine: (1) if they had received our letters, and if so (2) why they had not answered. Reasons for not responding ranged from "What letter?" to "I thought it was a joke!" to "I didn't bother to read the letter!" Eventually all 124 teachers replied to the invitation.

To us, the response was most gratifying. Only six teachers refused because they were not interested. Seventy-two agreed to participate according to the conditions specified, while 46 were unable to accept because of prior commitments. The nature and frequency of these prior commitments are listed below:

A.	Continuation of work on Master's degree in summer school	11
B.	No longer teaching physics	11
C.	Summer job commitment	10
D.	Physics not offered in 1967-68 at their school	5
E.	Health reasons prevent extended travel	3
F.	Misc. (changing jobs, expecting baby, etc.)	5
		46

In a number of instances when the teacher we attempted to contact was no longer teaching, a substitute teacher was offered by the school administration. However, we realized that accepting substitutes would not be consistent with randomization. The sine qua non of random selection is equal selection probability for all subjects. Teachers in the school system where a selected teacher no longer taught physics would have a higher selection probability than other teachers: a certain probability because of their own name on the master list and a certain additional probability because of being in the same school system as the selected but unavailable teacher. If substitutions were accepted, there would be a selection bias in favor of school systems with high turnover.

We are concerned about bias that might be introduced in the sample, so several characteristics of the acceptors

were compared with those who were unable to accept. A two tailed t-test was used to test the statistical significance of differences. These comparisons are summarized in Table 1.

[Insert Table 1 about here.]

Questionnaires were returned during 1966-67 from 117 teachers--71 acceptors and 46 non-acceptors. Apparently, teachers who accepted the invitation, when compared to the non-acceptors, are more likely to teach in larger schools and to be currently teaching the Physical Science Study Committee (PSSC) physics course.[2] It seems reasonable to interpret these differences as a greater receptiveness to innovation inlarger schools where previous innovations have been accepted.

A table of random numbers was used to assign 46 of the teachers to the experimental group and 26 to the control group. Letters describing the six-weeks summer session at Wellesley College were mailed to the experimental group together with letters of agreement to be signed by school officials. The control group teachers were invited to attend a two-day briefing session in Cambridge at which time their role in the evaluation was described.

[2]PSSC is a recently-developed physics course--one of the first of the national curriculum projects.

In the four months that followed the assignment to experimental and control groups, attrition occurred in both groups. Among the 46 teachers of the experimental group, one died suddenly, three quit teaching, one went to teach in Uganda, two were transferred to new positions, two could not come to terms with their school boards, and one became too ill to travel. Thus the experimental group was reduced to 36, or 78% of the initial group. We decided that this number was still adequate for our evaluation, and to avoid introducing bias we did not supplement the group in any of several ways that were considered. Five teachers in the control group dropped out for reasons similar to those above, leaving a total of 21 teachers (81%). Again this number was considered adequate for the research planned.

In 1966 similar attrition among volunteers occurred prior to a Project Physics summer briefing session. In that case, a sample of 45 teachers was selected from among volunteers to attend a summer session and then teach Project Physics. However, only 37, or 82%, actually attended the institute. Thus we are inclined to believe that the attrition of 20% over a few months is not peculiar to national random samples.

Our final sample consisted of 57 physics teachers. Thirty-six of these attended the summer training session

and are currently teaching the course. Another group of 21 attended the two-day briefing session and are presently teaching their regular physics courses. Data were gathered using a system of randomized data collection within each class which maximizes the number of tests and minimizes testing time for individual students (Walberg and Welch, 1967). Data are now being collected on the 3000 students being taught by the final sample of 57 teachers.[3]

Discussion

To claim a random sample, the population over which the sample is random must be specified. Our population is not, to begin with, "all high school physics teachers" but rather "all high school physics teachers on the 1966 NSTA list." There is undoubtedly some bias introduced by the estimated 20% of high school teachers who were not included in the list. Another bias is introduced by the exclusion of the 37% of the initial sample which had previous

[3]In addition to the 57 new teachers selected at random, 21 teachers experienced in teaching Project Physics are involved in the research design. A total of 4,067 students are being tested on approximately 80 different variables.

commitments. Since the prior commitment was often a summer institute we might infer that our final sample was biased toward the less able teacher. There were a substantial number of withdrawals after the assignment to groups. Although there is no clear distinction between the experimental and control, we must infer that the overall sample is again biased. We could characterize the population, then, as all high school physics teachers on the 1966 NSTA list who did not make summer commitments more than 3 months in advance and who would not back out of the research agreement. Certainly this population is more limited than would be ideal, but it is the best that we could possibly get and as far as we know it is better than any that has been used before.

Added to the possible biasing factors is the extra cost involved in random sampling. Purchasing the list of teachers, extra mailing costs, telephone calls, selecting the names at random, and many other clerical problems increased the cost approximately $1000 above what it would cost to use volunteers. The cost of bringing the 21 control group teachers to Cambridge for two days was $3500, an expense that could have been avoided if volunteer teachers were used. Further, travel costs of $5000 were required to bring the experimental group to the six-weeks session. This is approximately $3000 more than would be required for a regional sample. Thus the

total increase in cost by using a national random sample rather than a volunteer regional sample was an estimated $8000.

In general, we believe that our national sampling was successful. The teachers are cooperative and willing to assist in every way we have asked. The experience of working with them during the summer was rewarding and informative to the physicists and educators developing the course. We have been able to obtain information from a national sample of physics teachers. With this data we are able to make some generalizations to a national population on educational experience, training, teaching situation, knowledge of physics, and attitude towards teaching. Also, possible differences between these teachers and the volunteer teachers selected the preceeding year (Walberg and Welch, 1967) can be identified. These data should be helpful in judging the values and disadvantages of using volunteers for course development, a question which bears upon virtually all past evaluation of curricula designed for national usage. To the extent that the selected teachers have a representative sample of students, it will be possible to draw inferences concerning the attitudes, interests, and ability of students enrolled in high school physics. Most important, we are now able to try a new course

project in a variety of schools that are representative of all secondary schools.

An unexpected result has been the desire of those involved in writing the course to want to try out new ideas and materials on students of the randomly selected teachers rather than on the students of volunteer teachers. However, we have resisted the urge to try new things on these classes for fear of disturbing the experiment. Aside from the statistical reasons for taking national random samples, it would seem advisable to use such samples for subjective evaluation of course materials in their early stages of development.

We cannot say with certainty at this time that the invited teachers (experimental and control) will be consistent in their cooperation throughout the year. However, our experience with them thus far has given us little cause to worry. We hope to report at the conclusion of the evaluation that the random sample was worth the expense and effort.

Table 1

Comparisons Between Acceptors and Non-Acceptors of the Experimental Evaluation of Project Physics[1]

Variable	Acceptor Mean	SD	Non-Acceptor Mean	SD	Difference	t value
Grade 10-12 Enrollment	786	753	442	449	334	2.77*
Number of Physics Sections	2.3	1.7	2.0	3.5	.31	.65
Physics Enrollment	53	54	39	86	14	1.03
Fraction of 12th Graders Taking Physics	.28	.25	.26	.24	.02	.33
Fraction of Physics Students Taking PSSC Physics	.43	.75	.16	.56	.27	2.09*
Fraction of Girls in Physics	.15	.15	.20	.57	-.05	-.70

*p<.05

[1]The Non-Acceptor group includes the 46 teachers unable to participate because of prior commitment and the 6 who refused due to lack of interest.

References

National Science Teachers Association Staff Report. Some statistics of U.S. secondary schools, 1964-65. The Science Teacher, 1965, 32.

Walberg, Herbert J. and Welch, Wayne W. A new use of randomization in experimental curriculum evaluation. School Review, 1967, winter.

Walberg, Herbert J. and Welch, Wayne W. Personality and characteristics of innovative physics teachers. Journal of Creative Behavior, 1967, 1.

Welch, Wayne W. High school physics enrollments. Physics Today, 1967, 20, 9-13.

Welch, Wayne W. and Walberg, Herbert J. A design for curriculum evaluation. Unpublished manuscript. Harvard University.