

## WHEN IS IT RESEARCH?<sup>1</sup>

Annie W. Ward  
Volusia County

The writer, as a member of a hybrid profession- educational psychology - has frequently heard one branch engage in severe criticism of the research of the other branch. Incidentally, the criticism has been reciprocated. While it would be interesting to examine the merits of the contentions of both sides, it is the writer's position that concern for the improvement of educational research is more important than adjudicating the arguments. It is essential that those concerned about the improvement of educational research recognize forthrightly that research in education involves many more difficulties than research in other fields. This is true of even the other social sciences which, like education, are confronted by problems of the variability of the object being studied.- the human organism.

Attention to the appropriate and precise use of terminology is a logical first step in the improvement of educational research. The proper and precise use of terminology will clarify for the investigator the essential character of the project at hand and improve his communication with others who may be concerned with the progress of his work or who may become consumers of his findings. For instance, "research" is often used when such terms as "survey," "investigation" or "inquiry" would be far more appropriate and precise. Even greater liberties are taken with the term "experimental research," especially by those fairly skilled in statistical method when presenting objective data of some kind. All too frequently, those who use the term "experimental" have violated all, or most, of the fundamental rules of experimentation. The investigator of educational problems is seldom able to control the variables involved to the extent that "experimentation" requires. He usually cannot select individual pupils to serve as his S's or assign them randomly to "conditions." He usually cannot select the teachers who often must serve as his E's. Quite to the contrary, the pupils generally have been assigned to functioning classes on bases which are irrelevant to, if not inappropriate for, his study and the teachers are packaged with them by virtue of their prior employment and assignment. Studies conducted under such circumstances cannot be clear cut experimentation, analogous, for instance to those involving the learning of nonsense syllables under varied conditions of re-enforcement.

---

<sup>1</sup>The author presented the content of this article as an address to the 1960 meeting of the Research Section of the Florida Education Association and upon the Editor's request permitted its publication with minor editorial changes. -Editor..

However, this is not the type of information educators generally seek. Few in the public school are trying to conduct basic or pure research. The kind of research they do more properly should be designated "evaluative." Much, also, could be termed "action research" because there is group participation in planning, execution and evaluation. Typical undertakings include the preparation of local norms, comparison of achievement from year to year, identification of weak areas in the local instructional program, evaluation of new programs and of efforts to improve old ones, and local prediction studies. In all such studies, the principal concern is with situations in which change is continuously taking place - not only changes being measured, but other changes which may affect the results and of which the investigator is not even aware. Breadth of generalization is important, because the investigator wished to apply his findings in a wide variety of school situations. For this reason the limitations cannot be so rigorous that the results are not generally helpful. Such broad, evaluative research serves a professionally valuable purpose. Persons engaging in it should concentrate upon the means of improving it rather than wasting their energy on attempting to justify it as "experimental."

An incident from another discipline will serve to illustrate some of the pitfalls of evaluative research. Eysenck (2) reviewed several studies evaluating the effectiveness of psychotherapy and reported almost unanimous agreement that two out of three cases showed improvement or cure. However, the same percentage among the "untreated" showed remission of symptoms. Eysenck's statements have been quoted to demonstrate the necessity of a control group in evaluation studies. A close look at these studies having control groups points up some additional problems. In the first place, the "control" group in many studies was made up of those who applied for psychotherapy but dropped out either before treatment or after a few treatment sessions. There is therefore some question as to whether the controls can be considered as truly "untreated". Nor is there any evidence that the so called "control" groups were from the same population as the "treated" group. It is therefore very likely that the "controls" differed from the "treated" on many variables that were related to the remission of the symptoms.

The failure to provide a control group is probably the most common error of the novice in research. The logical necessity of doing so is so apparent that one is struck by the fact that many reputable universities have failed to teach their graduate students to design research studies incorporating them. Another all too common error is the failure to obtain pre-test data before starting treatment. There are certain principles of design that are so fundamental that they should be well and carefully taught before any university permits its graduate students to undertake research. It seems probable that graduate schools have improved their research training programs in recent years, but not all teachers holding the Master's degree today may be presumed to have mastered the most rudimentary principles of design. Ignorance of these principles on the part of administrators and supervisors may also be encountered, and because of their status in the power structure, "research" that is not in any sense "research" may be instituted.

Even when provision is made for a control group, it is very probable that it is not like the experimental group in some important way. That which is an important characteristic of one group, but not of the other, or unrecognized characteristics of both groups, all too frequently become uncontrolled variables which bias, if not vitiate, the data obtained and the conclusions drawn. This points up two needs which seem of greatest importance: (1) careful pre-evaluative studies to identify the most important variables operating in the situation under study; and (2) research designs which will make possible the control of the relevant variables. The first need often may be met by wide reading of the research literature in the area under study, because the common problems have been studied before. When adequate analyses have not been made previously, each investigator must conduct his own pre-evaluative studies to determine the functioning variables that will most probably be operative. The second need may be met in a number of ways, the most common of which is provision in the design for equalization, for counter-balancing, for randomization, for replication or for appropriate combinations of these techniques.

Equalization may be achieved by either of two general methods. Subjects in the experimental and in the control group may be matched on a one-to-one basis. The groups may be matched as groups on one or more of the most important variables and differences associated with other relevant variables may be controlled by the analysis of covariance and partial correlation.

The matching of individual subjects is generally very difficult. It is seldom possible to match many individuals exactly on more than one variable and almost impossible to find two classrooms in which the individuals can be matched on a one-to-one basis on every pertinent variable. The alternative is to match on only one variable and let the other variables "flow free" thus introducing unknown amounts of error. In large school systems it is often practicable to select sectional groups that are so similar with respect to all important variables that they may be considered to be matched. Under certain conditions it may be best to select from a large number of groups a sub-population of individuals who match on the most important variable and who do not differ significantly as groups on the other identified variables. When it is not possible or practicable to match on an individual or group basis, the analysis of covariance or partial correlation may be used to control existing differences in the uncontrolled variables.

Counterbalancing is often the simplest technique of control. Essentially each subject serves as his own control by being subjected to each treatment in rotation. Since the order in which the subject receives the treatments may influence the results obtained under each treatment, the design must provide for a "rotation" of the "rotations" of the treatments. For example, if three different treatments (teaching methods) were under study,

six different groups (classes) would be needed to completely counterbalance. This is true because there are six possible permutations of three things taken three at a time  $[k(k-1)]$ . Having six classes, each of which is to be subjected to each of three teaching methods - A, B, and C - an order of treatment like the following will permit the lumping of all treatment A, all of B, and all of C since all subjects will have been included in each and the order of presentation counter-balanced.

Table 1  
Schematic Presentation of Counterbalancing  
as a Control Technique

Class	Order of Treatment		
	First	Second	Third
1	A	B	C
2	B	A	C
3	C	B	A
4	A	C	B
5	B	C	A
6	C	A	B

Randomization is seldom possible in the school situation. The random assignment of children to sections and teachers to classes is a practical impossibility. It is often possible to randomize by class section if a large number of sectioned classes are available. However, replication can be used more often by building it into the design. By so doing the investigator will be relieved of the necessity of creating and maintaining the fiction that his data are derived from "random samples."

The importance and the practicality of replication merits emphasis. The investigator can build replication into his design more easily than he may realize and the results obtained thereby are much more generalizable than those obtained by most other treatments. An investigator who has twenty experimental classrooms matched with twenty control rooms can easily split both groups into ten classrooms each and treat each half as a separate experiment. If he has one hundred matched cases, or one hundred self controls, he can divide his subjects into two or more groups before data analysis and thus provide replication. The superiority of replicated experiments over the single experiment is clearly demonstrated in most standard references on research design. In fact, the replicated experimental design is logically more proper than the mass treatment in studies involving self contained units, such as the classroom, especially when the generalization to be made is to other self contained units possessing varying individuality.

A design that leads the investigator to work with "small numbers" in each of several experiments instead of "large numbers" in a single study may shock those who were indoctrinated into the worship of large numbers. It should be emphasized that there is a very important branch of statistics that deals with small sample techniques, and with the treatment of data that do not meet the rigorous assumptions underlying "classical statistics." In further support of studies of small groups, we have four important principles: (a) To be truly generalizable to individual cases, results must be consistent for almost every individual; (b) Small sample statistics which allow less room for individual variation are more appropriate for studies concerned with individuals, or with self-contained units possessing an individuality, than are large group treatments which blur the very thing being studied; (c) Many replications made possible by the use of small samples lead to more confidence in the results than will a single replication with a large sample; and (d) The nature of the data collected on human organisms does not justify the elaborate procedures of many classical treatments. For further treatment of these points, the reader is referred to Siegel (3) and Tate (4).

This paper has not really answered the question "When is it Research?" It is simple to take a negative approach, merely pointing up the things that so weaken the findings that by any reasonable standards, they cannot be considered research. Rummel states that "research is an endeavor to discover, develop and verify knowledge (2, p. 2)." He also points out that "the progress of man in solving problems . . . . has been concomitant with the discovery and development of methodological procedures . . . . (which make) . . . . the research process itself more rigorous, more discriminating and more dependable (2, p. 2)." It is hoped that these suggestions will help to make progress in educational problem-solving activities, particularly those involving evaluative research:

1. Concentrate on the problem. Have some well defined objectives and make the "statement of the problem" clearly specify the job to be done.
2. Determine the most appropriate method of attacking the problem. A formal study is not always necessary. If a study is to be conducted it is absolutely necessary to decide upon the criterion variables and upon the method of measuring these variables.
3. Identify relevant variables, i. e., those which may affect the criterion variables, and which must, therefore, be controlled, and include the control method in the research design.
4. In reporting the study, indicate clearly what variables were controlled, and how, so others may know the extent to which the results of your study and generalizable to their own situation.

## Bibliography

1. Eysenck, H.J. "The effects of psychotherapy: An Evaluation," Journal of Consulting Psychology, 1952, 16, 319-323.
2. Rummel, J. Francis, An introduction to research procedures in education, New York: Harper and Brothers, 1958.
3. Siegel, Sidney, Nonparametric statistics for the behavioral sciences. McGraw-Hill Book Company, Inc., 1956.
4. Tate, Merle W, and Clelland, Richard C., Nonparametric and shortcut statistics in the social, biological and medical sciences. Danville, Illinois: Interstate Printers and Publishers, 1957.