DOCUMENT RESUME

ED 030 209 EA 002 351

By - Trismen, Donald A.

Management and Generalizability.

Pub Date 8 Feb 69

Note - 7p.; Paper presented at the Annual Meeting of the Amer. Educ. Res. Assn. (Los Angeles, Calif., Feb 8, 1969).

EDRS Price MF - $0.25 HC - $0.45

Descriptors - *Educational Research, *Educational Researchers, Evaluation Criteria, *Information Systems, *Research Problems, *Research Tools

Although management of a large-scale research study is the responsibility of the team of investigators, control of the educational treatment is often shared with teachers and administrators in participating schools. This results in continuous modification of objectives and procedures which, in the traditional view, is poor scientific practice and prevents generalizability. To compensate for the lack of control, (1) a broad range of data from very large populations should be collected to allow analysis of all major variables, and (2) an educational information system should be created to provide the information necessary to evaluate replications of treatments across schools. This approach differs from a traditional research and evaluation study in determining under what conditions a particular effect occurred rather than trying to specify and enforce a particular set of conditions. (MF)

Preface

Naming a child or a paper at birth is a hazardous undertaking. As the child or the paper grows, gradually acquiring a personality of its own, there is always the danger that the chosen name will be completely inappropriate to the characteristics of the adult. With respect to my children, it's too soon to judge. In the case of this paper, it grew in some rather surprising directions. Luckily, it's somewhat easier to rename a paper than a child, and this one is now called Management and Generalizability.

MANAGEMENT AND GENERALIZABILITY*

Donald A. Trismen
Educational Studies Division
Educational Testing Service
Princeton, N. J.

For the purpose of considering management problems, it is useful to make a distinction between managing the study and managing the situation to be studied. Managing the study raises concerns about procedures, materials, and personnel. It involves the development and production of instruments, selection and orientation of field representatives, sample selection, gathering and processing of data, and coordination of all these activities in a manner consistent with the overall goals of the research. In large-scale research, the problems of managing the study are complex and worthy of attention in their own right. However, the following remarks will be concerned only with management of the situation to be studied.

These two aspects of management raise quite different questions. In managing the study, the question is "how best to manage." It is assumed that the more efficient and effective the management, the better. Management of the situation to be studied is often called experimental control. The relevant question here is "how much, or even whether, to manage." Perhaps in attempting to answer this question, the first thing to consider is the feasibility of experimental control in large-scale research. In educational research on any scale, subjects are exposed to some more or less specified treatment, and measurements are taken to

---

*Prepared for a symposium entitled "Management Problems in Conducting Large-Scale Research and Evaluation Studies," held February 8, 1969, at the Annual Meeting of the American Educational Research Association, in Los Angeles.

describe the effects of this treatment on these subjects. The classical research paradigm requires that the treatment be controlled (or managed), but that the subjects be free to respond independently. In small-scale studies, this control is exercised by the investigator alone or by a small group of associates responsible directly to him. In addition, the investigators perform the important function of monitoring the treatment to assure its uniformity for all subjects--a kind of quality control.

The large-scale research study, as a consequence of its size, introduces new elements into the situation just described. Although the management of the study per se remains a responsibility of the team of investigators, the control of the educational treatment is often shared with teachers and administrators in participating schools. The investigators' control over the experimental situation has thus been weakened by the introduction of several intermediaries. Although the research paradigm includes these additional human elements as part of the treatment, they are not under the control of the investigator in the way that the inanimate treatment components are. Not only are teachers and administrators the agents through which the treatment is applied, but they are also subjects in the sense that they too are affected by the treatment. Unlike the investigator, they are student-oriented rather than treatment-oriented. As first-hand observers of the experimental situation, they have available to them numerous information feedback channels other than the data of the study. They are not inclined to await the outcome of a formal analysis of the data before evaluating treatments and adapting them to the needs of individual students. They find it incompatible with their training and experience to interact with each student identically and independently. One would

have difficulty finding a group less suited to the task of applying an educational treatment in the objective manner required by the traditional research model.

Thus a realistic look at the typical field situation forces us to admit something which has been obvious to educators for a long time--one of the most important effects an educational treatment can have is the effect upon its own characteristics, through the agency of the participants. In other words, treatment characteristics can be considered as a dependent variable to be described during the course of a study, as well as an independent variable to be specified at the beginning. The traditional research model requires that the treatment be specified and controlled throughout, but this runs counter to the pedagogical instincts of teachers and is nearly impossible to achieve in a large-scale research study.

The characteristics of multiple programs, commonly the focus of large-scale evaluation efforts, are usually the despair of the educational researcher. Overlap of students, teachers, objectives, and criteria among programs, self-selection of experimental groups, and continuous modification of objectives and procedures are all poor scientific practice, at least in the traditional view. However, they can also be feasible and adaptive modes of behavior in the real educational world.

Assuming that control of the experimental situation cannot or will not be adequately performed in a large-scale educational research study, what are the consequences? For one thing, it is highly likely we will end up studying not only the anticipated treatment, but also many variations of it bearing the same name. Thus, in addition to a wide

variety of criterion data, there must also exist a broad range of data descriptive of the treatment itself and the environment in which it is applied. In other words, when the opportunity is not available to specify the treatment, then the obligation to describe it after the fact becomes even more crucial. Of course, any generalization of our results must now be made to treatments and environments similar to those which evolved during the course of the study. Moreover, an essential feature of these treatments and environments is the fact that they were not under control of the investigator during the original study. We can only hope for replication of results when we start with subjects, treatments, and environments similar to those of the original study and when the experimental conditions evolve in ways similar to those of the original study. This is rather a lot to hope for, and perhaps explains some of the difficulty encountered in replicating educational studies.

However, the difficulties of exact replication should not deter us from determining what in fact does happen when the conditions of replication vary in known and unknown ways from the original conditions. Given our relative lack of control over the conditions of education in most situations, one could argue that it is the robustness of educational treatments which is of greatest practical importance. Perhaps like automobiles and home appliances educational treatments should be designed and tested for satisfactory performance under widely varying conditions of misuse.

What are the implications of replication for large-scale research and evaluation studies? First, emphasis must shift away from the one-shot study. Instead, there must be widespread, continuing collection of

a broad range of data in a large number of naturalistic settings. Only in very large populations will the interactions of interest among student, treatment, and environmental variables be represented in numbers of cases sufficient for appropriate analysis. Emphasis will be placed on hypothesis generation, rather than hypothesis testing.

Second, if results across replications are to be compared, there must exist at least a common core of data. However, because of desirable (or unavoidable) variation among the various conditions of replication, there will also be a demand for data unique to local situations. Clearly indicated is the need for an integrated network of data collection, analysis, and reporting procedures--an educational information system. Such a system would provide the data necessary to evaluate replications of treatments across schools. In addition, the data would serve a monitoring and feedback function for each individual school. The situation to be monitored would be one of the school's own choosing, rather than one imposed from without and implemented by unenthusiastic participants within. Analyses and reports would present and interpret information both across and within schools. Treatments could be altered at any time by participating schools, as continuously updated feedback data indicated the need.

How are the management problems of such an information system different from those of a traditional research and evaluation study? First, the efforts of the system managers are directed toward determining under what conditions a particular effect occurred, rather than trying to specify and enforce a particular set of conditions. This approach requires extensive data collection, especially via questionnaires and observational instruments. On the other hand, it requires no control of

the situation. Second, the data are more apt to be used for local decisions which must be made in real time. "Should the foreign language program be offered at the sixth grade level next year?" "Which students should be admitted to the academic track?" Pressure for the acquisition, processing, and reporting of data in time to have impact on decisions such as these requires adherance to data processing schedules far more demanding than those of the typical research study. If the educational researcher hopes to see his work have real influence on educational practice, he must be willing to pay the price of relinquishing the unhurried pace of laboratory research. If he expects to create a demand for his product, then he must accept responsibility for producing the supply when it is needed.

In summary, large-scale replication under a variety of conditions serves as a means of mounting an inelegant, broadside attack upon the problem of demonstrating the effectiveness of educational treatments. From the standpoint of research efficiency, it is a giant step backward from the well-controlled "crucial" experiment, designed to enable a precise statement of generalizability. However, in the large-scale educational research study, such control is more often a wish than a reality. Perhaps the time has come to substitute descriptions of what <u>is</u> happening in education for statements of what <u>would</u> happen <u>if</u>.