METHODOLOGICAL HAZARDS IN

CONDUCTING EVALUATION RESEARCH

Mark Sizemore
Department of Sociology
North Texas State University

Marvin Ernst, Ph.D.
Dallas Geriatric
Research Institute

April 1978

Presented at the Southwest Social Science Convention, April 12-15, Houston, Texas.

Introduction

Academic researchers have begun to pay attention to the promises and problems of evaluation research. In the former case, researchers have begun to realize that evaluation research is on the verge of gaining considerable respect as an area of research endeavor. It offers, for instance, the opportunity for the researcher to become actively involved in testing out service delivery concepts and brings the sociologist face-to-face with the problems of the "real" world. Evaluation research monies have also become available during a time of dwindling support for the more esoteric types of research, thus affording the opportunity for funding of various kinds of data collection. Evaluation research also is conducted at the intersect of theory and practice. Within this confluence, the researcher often finds himself/herself in an exciting area of research where problems of human and agency performance are dealt with in a practical manner.

Unfortunately, while evaluation research has many promises, it is often frought with many potential pitfalls for the researcher. Beyond the difficulties of negotiating the research and the politics of evaluation, a number of hazards may seriously impede both the progress of the evaluation effort and the decisions made from the data.

These hazards are of such a nature that failure to recognize them from the beginning of a project can lead to considerable frustration on the part of the agency being evaluated, and the researcher(s) who are conducting the project.

This paper is based on direct experience in a number of evaluation projects. Materials delineated have been selected by the authors to represent major concerns which need to be dealt with in any evaluation effort. These problem areas include: 1) the definition of the variables to be assessed; 2) the delineation of the research population; 3) the process of sample selection; 4) the implementation of the research in the field, and 5) the statistical presentation of the data.

Definition of the Variables

Considerable difficulty may arise in the process of clarifying what is to be measured. Often a semantic problem, the agency to be evaluated and the evaluator may talk past each other when discussing exactly what variables are to be measured. This may include terms which have specialized meaning to each of the two actors. A sociologist, for instance, may use an abstract term which he/she feels is fairly obvious to other people and may have the meaning misinterpreted by the agency. A practitioner, on-the-other-hand, may use jargon which is in vogue within the agency, but which may not be tied to other conceptual frameworks. An example of this may be the word 'dependency.' The practitioner might be thinking in terms of a specific instance of a dependent state, such as dependency upon the family, while the evaluation researcher is conceptualizing dependency as a role which the individual increasingly adopts with advancing age (Atchley, 1972: 103), and which may be internally or externally motivated. The researcher may also be considering dependency as partly a function of a culturally bound social system to which elements of independence or dependence may be allowed. The important point is simply that the researcher may draw the impression that the exact variables for measurement have been agreed upon, when the practitioner simply wants a yes or no answer as to whether his/her workers are fostering independence among the elderly.

Clarification may also be needed around the issue of program operations versus program impact. While the evaluation researcher is specifying impact, the effect of the overall program upon the life condition of selected participants, the practitioner may be talking about program operations, or how effective or ineffective a given set of procedures happens to be. While impact evaluation is assumed to be somewhat dependent upon the procedures of operation, the decisions reached at the end may be at variance, for the same outcome may be the result of many different procedures.

Consequently, the evaluator must clarify with the agency the focus of the evaluation effort. He/she must also recognize that the direction of the study must flow out of the agency needs, rather than one's own professional interest. Most sociological researchers would be more interested in the impact of programs because of the opportunity to study relationships between variables than in studying whether procedure A or B got more services to the client.

Since agreement on the definition of what is to be studied precedes actual measurement, it behoves the researcher to clarify initially all relevant variables to be included.

Definition of the Population

A restriction imposed upon the sociological evaluator is that many agencies simply will not be able to fund data collection on individuals who are outside of the service structure which they provide. Or, because of political reasons, they may be unwilling to fund these kinds of data collection. This restricts the researcher's ability to define a broad based population for comparative purposes. Difficulties inference and generalization are inherent in a limited population.

An additional problem lies in the fact that most agency populations are fluid. If the evaluation is to be studied over time, the population base may not be the same at Time One and at Time Two. This could be caused by changes in intake procedures or new legislative and bureaucratic mandates. Intake procedures may be limited or expanded according to funding availability, thus expanding and constricting the population. What does the researcher do, for instance, when midway through the evaluation project it is learned that one-third of the inital sample is going to be dropped from services due to alterations in the funding requirements?

While the evaluation researcher will never be able to affect control over the implementation of specific programs, he or she can be sensitive to the changing nature of the population base and design the evaluation with these factors in mind, or can attempt to impress upon the agency the need for concern regarding consistency within the population to be studied.

Sample Selection

Theoretically, the researcher develops the research design in such a manner as to be appropriate for obtaining data necessary to the decision-making process. Although classical experimental design is attractive, utilization of quasi-experimental design (Campbell and Stanley, 1963: 171-246) is frequently the most practical alternative. This may require the selection of varied subtypes of samples within the overall population base. The definition of these subtypes must be agreed upon by the researcher and the agency being evaluated. At this point, there is considerable need for the researcher to maintain a conceptual purity regarding the mix of service types within any particular subgroup. Frequently, service deliverers define subgroups on the basis of services provided. This tends to mask the commonalities between subgroups who may be very similar in nature and who receive services via a similar mechanism. For example, the key analytical bases to a subgroup may be the intimacy of service rather than the fact that one receives some kind of homemaker service by a specific individual while another person receives meals-on-wheels by a specific person. The commonality for the subtype is the intimacy and frequency of interaction rather than the specific category of service.

The sampling technique may also be heavily influenced by the budgetary time frame of the agency, rather than the requirement of the research design. Agencies are frequently under a time press to get the data in on which to make a decision. The researcher must be cognizant of this fact when preparing measurement assessment.

A number of measures can discern changes over time, but the time frame must be of such a length that change can be legitimately expected. An example of this is the Havighurst Life Satisfaction Index A. Life satisfaction is based upon the individual's perception of previous life situations. These are not likely to vary widely during a three to six month period.

A third problem which occurs during the sampling phase lies in the fact that the evaluation effort may be agreed upon at one level of each agency and implemented at another level. Poor vertical communication may result in an erroneous perception of the evaluation. Since agency administrators and service personnel are engaged in providing direct service to the client, lack of communication may seriously hamper the researcher's effort to draw a representative sample.

Since lower level staff are never presented a coherent picture of the research process and design, they may seriously influence sample selection by failure to adequately support the research staff. Client data may not be organized in such a manner as to facilitate sample selection. Key respondents may be told by staff members that they do not have to participate and that they should refuse to respond to the research staff.

In addition, the administrative structure may be such that the chief administrator suggests key procedures for sample selection which, in reality, are not functional. For example, the researcher and the administrator may agree that the best way to select the sample is on a random basis from client records. Unfortunately, it may happen that when the research staff attempts to draw the sample they find that the records are not up-dated frequently and that many new clients have not been recorded, or deceased clients never removed from the record file. This not only causes serious problems in drawing the sample, but also raises legitimate questions of the researcher's ability to generalize from the data to the entire client population.

Implementation

Not dissimilar in nature to other research efforts, gathering the data during an evaluation has several pitfalls. While survey methods of data collection have been criticized by a number of people, few alternatives are available for a community based population.

Typical problems of refusal rates and non-cooperative responses are compounded by the fact that the research is being conducted on a specific agency which provides a specific service. Clients may perceive the research staff as part of the service delivery agency. Since many of them fear losing services, they may mask many of their real responses in order to continue receiving the services.

It is also possible that workers who move from client to client bear with them tales regarding the interview, hence heavily influencing subsequent responses. This is particularly evident in cases where the service staff has not been informed of the nature of the interview, and are somewhat afraid of what the results will say.

If the data collection deals with agency personnel rather than clients, considerable amounts of care must be taken to insure that the data is collected within a relatively short time frame. If interviews are strung out over an extended period, shop discussions may result in a homogenization of the data since employees have gotten together to discuss interview responses and may bias later interviews. This may result in the measurement of influencials rather than the true feelings of the overall organizational structure.

Data Presentation

When preparing to present the findings to the agency, the evaluator should consider the possibility of differences between the interest of the sociologist, on the one hand, and the role of evaluator on the other. Frequently, sociologists desire to look at the data in great depth. They desire to explore relationships and to test various hypotheses. The agency, however, is principally interested in making policy decisions. Statistics which can deal in depth with the statistical relationships may not be understood by the agency personnel. Over simplification of complex relationships by the agency personnel may result in the perception of a .05 probability as a finite conclusion, rather than recognizing the element of chance which exists.

Conversely, overly simplistic statistic presentation may result in erroneous interpretations of the true relationship. Eye-balling a bivariate table, for instance, can result in extremely hap-hazard data analysis.

Finally, the agency who is being evaluated is often under a considerably stricter time schedule. Demands for data on which to make decisions may be placed upon the researcher during which the researcher must avoid premature exposure of the data which subsequently must be revised after further data analysis.

Conclusion

The primary focus of the paper is to draw the attention of the sociologist engaged in evaluative research to several problematic areas. Although the literature abounds with descriptions of difficulties encountered in evaluation research, the foregoing is principally derived from direct experience of evaluations on service delivery agencies.

The sociologist, whose academic training has concentrated on "ideal" types of research, may not be prepared for dealing with problematic situations encountered during a program evaluation which does not quite conform to ideal patterns. In addition to wasting time and energy, the sociologist may also experience considerable frustration which may have a negative effect on morale.

Because of its nature, evaluation research has a number of pitfalls which the researcher must be aware of, even before contact between agencies is made. Since mental sets develops early in the negotiation process, a thorough working knowledge of potential hazards may assist the researcher in planning a course of action.

BIBLIOGRAPHY

Atchley, Robert C. The Social Forces in Later Life: An Introduction to Social Gerontology, Wadworth Pub. Co., Belmont, California; 1972.

Campbell, D.T. and J.C. Stanley. "Experimental and Quasi-experimental Designs for Research on Teaching," in Handbook of Research on Teaching, Ed. N.L. Gage (Chicago, Rand McNally & Co.); 1963.

Weiss, Carol H. Evaluation Research (Englewood Cliffs, N.J.: Prentice-Hall, Inc.) 1972.

END