

designing evaluative research

James Longest

A great potential for progress in evaluative research exists today because of the very nature of our society. As the demand increases for social change and problems arise in satisfying these demands, there's also a concern for assessing the effectiveness of social action programs. Educators, legislators, and businessmen are but a few who recognize the importance of sound evaluation for increasing efficiency. Probably no type of research is more in demand than evaluative research.

However, probably no type of research is so little understood and so seldom well executed. Some authors claim that the theory and method of evaluative research haven't developed to the same extent scientific methodology of nonevaluative research has.¹ One of the contributing factors to the confusion about evaluative research may be the prevailing belief that evaluative research is to be classified as "applied" rather than "basic" and doesn't require the best research methodology and interpretation within a well-defined theoretical framework.

Some claim that problems posed by evaluative research are too complex and can't be placed within a framework of experimental design. Campbell and Stanley have presented a more realistic interpretation of the general applicability of experimental design. They affirm that although experimental design is still ideal, in reality it's nothing more than stating what we want to attain by applying the scientific method. With this interpretation, evaluative research, which uses the scientific method to evaluate, fits into the framework of experimental design. The problems are researchable.

In some evaluative research, as in all research, it's difficult to achieve the scientific rigor desired. Until recently, many social scientists have tried to copy some physical scientists in trying to research only those problems that fit into a micro and single dependent variable design. Providing adequate answers to many of the questions asked by evaluative research problems necessitates analyzing multi-dependent as well as multi-indepen-

James Longest: Professor of Rural Sociology, Department of Agricultural Education and Extension, College of Agriculture, University of Maryland.

dent variables. The technology, statistical methods, and methodology exists, but must be further developed to be adequate.

In evaluative research, we should strive to achieve the highest possible degree of scientific methodology and theory testing for conducting and controlling valid and reliable evaluation. As Campbell and Stanley point out, "We must increase our time perspective, and recognize that continuous multiple experimentation is more typical of science than once and for all definitive experiments."²

Design Elements

The first element for designing evaluation research is: *define the general conditions and problems* that require an evaluation to be made. When possible, all aspects related to the problems should be specified.

The second element required is: *specification of theories, assumptions, and values* that are accepted as true and will guide the formulation of the evaluation research. For example, in poverty research, the theory of "hierarchy of needs" may seem applicable to the motivation of poor families. Therefore, this theory could be the basis for a number of hypotheses in evaluation research on poverty.

On the other hand, the hypothesis that boundary maintenance tends to inhibit coordination of agency contributions and that rigidity of organizational functions assures resistance to change might be used to guide research on why organizational resources haven't been able to adequately help poor families. It should be obvious that the theories, assumptions, and values assumed as relevant are crucial to determining *what is to be evaluated and what the hypotheses are*.

One of the contributing factors to the confusion about evaluative research may be the prevailing belief that evaluative research is to be classified as "applied" rather than "basic" and doesn't require the best research methodology and interpretation within a well-defined theoretical framework.

The third element to evaluative research design is: *an explicit statement of what is to be evaluated*. The question, "Is Program A effective in changing subjects from position P to position T?" is quite different from the question, "Is Program A more efficient than Program B in producing effect Y?"

Having stated clearly what is to be evaluated, the fourth essential element is: *formulation of the hypotheses, identification of the variables, and specification of the units of analysis*.

The fifth element is: *the how, when, and where for data collection*. The sampling procedure, when used, must be specified. The instruments and methods for obtaining data should be selected or developed.

Evaluation Design of Social Change Programs

The sixth major element for research design is: *how the data will be processed*. This includes plans for coding, organizing, and the general program format for processing the data.

The seventh element is: *analysis of data*, which is guided by the theories and hypotheses formulated earlier. The analytical techniques and statistical procedures need to be planned in detail.

The eighth element is: *plans for reporting the information*. This will include the format of the report and the intended audiences. If there is to be a preliminary report of findings prior to the final report, this should be noted too.

The ninth and last element of an evaluation design is: *a schedule for the execution* of the evaluative research. This includes a time schedule for the research, the kind of project personnel and other resource requirements in each of the stages, and the source for each of the resources. Much of the scheduling for evaluation has to do with the *how* and *when* of the evaluation—such as, what measures are to be taken, treatments given, in what order, and at what intervals.

Statements of *what is to be evaluated, how, and when* will logically suggest many of the characteristics of design. For example, if the objective is to determine how much more knowledge exists after as opposed to before treatment X, then a before-and-after research design is required. If the hypothesis is that treatment X is more effective than Y, then matched groups or equivalent random or statistical controls, with before-and-after measures for comparison can be used.

Each of these nine elements are necessary for designing sound evaluative research. Each fits into the total design. The strength or weakness of any one element can certainly contribute to the success or failure of the entire research project. As can be observed, the elements of evaluative research design are the same as those for any other well-designed research.

Since much of the confusion concerning evaluation research and its capabilities has stemmed from the recent demand for assessment of social action programs, let's focus on that type of evaluation.

Essentially, there are four levels at which evaluation of social action programs may be undertaken. These have been conceptualized by Stufflebeam³ of Ohio State University Evaluation Center as:

1. Context evaluation.
2. Input evaluation.
3. Process evaluation.
4. Product evaluation.

These different levels of evaluation are definitely interdependent. A thorough evaluation of any program can't take place unless

all four types of evaluation are executed and interrelated in the evaluation design.

**Context
Evaluation**

Context evaluation can be used to evaluate social action programs when undertaken in a number of different contexts. This level of evaluation often focuses on the *type of social structure* that includes the nature of the “power structure” and how it relates to the action program being carried out. Since environment can be a contributing factor for differences in program outcomes, context evaluation may also investigate the historical and cultural environment of the action program.

Context variables are most often considered independent or control variables, but can be used as intervening variables. When a particular action program environment is desired, context variables may serve as input or treatment variables to create the desired milieu.

In most evaluative research, a certain amount of context evaluation takes place as fundamental information for completion of the research design. In some cases, hypotheses concerning strategy and likely outcome may depend in great part on particular characteristics of the context. For example, action program results may be quite different in social systems where the power structure is controlled by a few persons compared with systems where there are multipower structures. Context evaluation is then necessary.

Context evaluation provides for a more objective assessment in *input*, *process*, and *product* levels of evaluation. The nature of the context will determine in part the amount and quality of required inputs, the process to be followed, and the product that can be expected with a given resource.

Here’s an example of the above principles. Suppose an agency wishes to establish a program to help “the elderly in poverty” in two adjacent counties in a state. The two counties each have about the same population, in the 60,000 to 75,000 range. The proportion of elderly in the two counties is in the range of 15% to 20% in both of them. County A is one tier of counties closer to a metropolitan center than is County B. Here’s a comparison of the two counties:

1. Both are structured differently—in County A they have 5 county commissioners and a county executive, while in County B they have only 3 county commissioners.
2. In County A, a higher proportion of the population has a high level of education, and there are generally more and higher quality schools including a community college than is characteristic of County B.
3. County A has organized groups for all major interest areas, including a Golden Age Club, while County B

appears less organized and has no organization representing the interests of the elderly.

4. People acquainted with both counties believe they're different in other undefined ways.

Thus, context evaluation is needed because apparent differences exist in (1) power or leadership structures, (2) valuing or support for education, and (3) organized groups to represent the elderly and other interest groups. These differences, the historical and cultural reasons for their existence, and the need to define what other differences and similarities exist require context evaluation. These differences indicate different approaches are needed to aid the elderly in each of the counties. Knowledge of the contexts is needed to interpret differences in community and older people's reactions to program inputs and processes. Context differences also call for adjustments in input and process procedures. Otherwise, very different product results could be obtained from having applied identical social action programs to these two areas.

Unexplained product differences from application of identical input and process procedures to a number of counties often remain unexplained because contexts were assumed equal when, in fact, they were very different.

Input Evaluation

Input evaluation investigates the action program resource requirements. Examples of such resource inputs are personnel, program activities, facilities, equipment, and needed training.⁴ Input evaluation can be used in creating new programs or in assessing ongoing programs. For new action programs, it poses the question, "What kind of resource inputs are needed?" Both input and context evaluation can be helpful in assessing whether resources are realistic in terms of the context, the task, and the terminal objectives.

When programs are in process, input evaluation asks the question, "How adequate are inputs for meeting program objectives within the present framework?"⁵ In this way, input and context evaluation can be useful indicators for needed program revision. This might involve changing inputs to accomplish the goals and objectives or changing objectives when inputs aren't adequate.

In our illustration on "the elderly in poverty" program discussed in context evaluation above, the context differences called for differences in inputs. One of those differences was that County A had an organization representing the interests of the aged, while County B didn't. Since governmental and other resources are rarely systematically applied to help the impoverished unless some type of influence is exerted by an organization or a coalition of organizations, County A is in a more advanced stage of readiness to act since they already have an organized group representing the elderly. More atten-

tion and effort will be needed in County B to identify and train leadership since no organization exists that would have identified, trained, and used the elderly.

A closely allied need is the identification of important organizations and influential citizens in the county's power structure. In this case, the elderly club in County A probably will have done some of the groundwork, while in County B most all of such supportive linkages will have to be established in forming and operating an effective organization. Hopefully, context evaluation of leadership structures in both counties will have identified influential organizations and individuals, but in County B more remains to be done to effectively use their inputs to aid the elderly poor. Once programs are launched, each county's progress would need to be evaluated for how adequate the inputs are for meeting the program objectives.

Process Evaluation

Process evaluation is one of the most time consuming and resource demanding. It's probably one of the most difficult to do well, and one of the most neglected types of evaluation.

Process evaluation investigates the *program operation . . . the process . . . the procedures*. Since process may be a crucial intervening variable between input and product, it must be assessed for complete program evaluation. Failure to include process evaluation might lead to inaccurate conclusions in program evaluation.

Good process evaluation requires the investigator to be on hand when crucial program elements are being executed. Otherwise, the evaluation may not be able to explain either a negative or an extremely positive result. It also enables the evaluator to record unplanned occurrences which could be significant to the outcome of the program.

As with input evaluation, process evaluation is an integral part of the program planning and execution process. Such evaluation can point out the in-process difficulties and recommend alternatives for overcoming obstacles or for improving the program.

Our "elderly in poverty" example offers an illustration of the use and importance of process evaluation. Suppose that an objective is to obtain food stamps for the elderly poor in both counties. Suppose also that it's decided that presenting facts and cases to illustrate food stamps are needed by many of the elderly will cause an effective food stamp program to be established by the governments and that all eligible elderly will be informed and encouraged to apply.

Suppose further that the approach is used in both counties, but that in County A the government doesn't act, while in County B it does. If process evaluation is being used, the reasons for failure in the one case and success in the other will probably be known because the process was monitored. If it weren't monitored, a post evaluation would have to be initiated to

determine the reasons for failure in A and success in B. A post analysis may be less effective, however, because facts of actual behavior and reasons for decisions may be concealed or otherwise unavailable.

With process evaluation, County A's problems might have been revealed before rejection and a change in inputs initiated to have altered the results. For example, suppose the five commissioners in County A had heard the facts from elderly citizens and program personnel and seemed favorable. Suppose further, however, that the county executive, because of misinformation from other sources, reacted negatively to the council's recommendations and the council then voted against the program at their next meeting. If process evaluation had been operative, the negative response from the executive should have been known and counter influences exerted on him and council members before their next meeting. The effect could well have meant approval rather than rejection of the program.

In any event, a systematic monitoring of process should provide knowledge on reasons for successes and failures and therefore knowledge on what corrective action is needed.

Product Evaluation

One of the most crucial types of evaluation is product evaluation. It focuses on *how well program objectives have been met*. It can expose strengths or weaknesses in context, input, and process variables.

Although it's one type of evaluation most often used, there are current feelings that product evaluation is too difficult and not even necessary. This is based on the idea that all that's needed is input and process evaluation. This is unfortunate, for often input and process evaluation are unable to conclude whether goals are realistic given the time schedule and other resource allotments. Often these factors can't be fully evaluated until the action program is completed. This requires product evaluation.

Product evaluation can be useful in determining what kind of changes are necessary in context, input, and process variables to achieve product quality as well as quantity.

In the example of the elderly, assume that the food stamp programs were eventually established in both counties. Assume further that context, input, and process evaluation indicated that the programs were succeeding in both counties. A doubting Thomas suggested, however, that it might be interesting to check by actually finding out how many of the elderly in need of food stamps actually (1) were informed of their eligibility and (2) had applied for them successfully if they so desired. In other words, he was suggesting product evaluation. Suppose further that product evaluation was done, and in County B of all those eligible, 80% had been informed and 70% of those who were interested had enrolled successfully. In contrast, in

County A only 50% had been informed, and of those informed and desiring aid only 20% had actually enrolled.

This, then, would be a case where early in a program, product evaluation would discover a large difference in results that weren't evident by use of the other types of evaluation. This discovery should then lead to more in-depth context, input, and process evaluation analyses to discover why results were so low in County A.

Conclusion

Some of society's most debilitating, costly, and expanding social problems require evaluative research within the best known applicable research designs. Social science theory and knowledge are advanced enough and the technology does exist to feasibly execute basic evaluative research. We need only be creative enough to develop and design the instruments and procedures.

Footnotes

1. Edward A. Suchman, *Evaluative Research* (New York: Russell Sage Foundation, 1967), p. 6.
2. Donald T. Campbell and J. C. Stanley, *Experimental and Quasi-Experimental Designs for Research* (Chicago: Rand McNally & Co., 1966), p. 3.
3. Daniel L. Stufflebeam, "Toward a Science of Educational Evaluation," *Educational Technology*, VIII (July, 1968), 5-12.
4. *Ibid.*
5. *Ibid.*