DOCUMENT RESUME

ED 073 370 AC 014 253

AUTHOR	Bennett, Claude F.
TITLE	A Hierarchy of Program Evaluation Applied to Volunteer Leadership Development.
INSTITUTION	Federal Extension Service (DOA), Washington, D.C.
REPORT NO	PSD-1-6-2-72
PUB DATE	Feb 72
NOTE	14p.; Revision of presentation to 4-H Youth Development Workshop on "Utilization of Current Knowledge about Voluntary Participation, Evaluation and Application to Problem Priorities," National 4-H Center, Feb. 14-18, 1972
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Evaluation Criteria; *Evaluation Methods; Extension Education; *Leadership Training; *Program Evaluation; Speeches; *Volunteer Training; Youth Clubs; *Youth Leaders

ABSTRACT

Plans for evaluation should be laid prior to implementation of a program. A hierarchy of types of program evaluation helps select the exact type of evaluation intended. This hierarchy starts with the easiest (and least useful) type of evaluation and six progressively more difficult levels of evaluation are added in cumulative fashion. The seventh and final type is a master plan for evaluation, which should be accomplished by starting with the top steps and their interrelationships and working down. The evaluation types, with applicable questions for extension leader orientation, are: (1) inputs made--How much time and money are expended in leader orientation, and is this the right amount? (2) activities performed--What kinds of orientation activities are new leaders engaged in, and are there the right number of activities? (3) recipients involved--With how many leaders do we allocate certain amounts of orientation time and activity? (4) reactions--How do volunteer leaders react to their orientation? (5) "KAS" change (change in knowledge, attitudes, and skills)--Which development are a consequence of orientation? (6) "practice" change--Do the attitudes, skills, and knowledge gained through orientation help the volunteer in working with 4-H participants? (7) results achieved--Do the volunteer roles help achieve ultimate aims with 4-H youth? Evaluation at the lower levels is meaningless without answers to higher evaluative questions. (KM)

Extension Service, USDA
Program and Staff Development

A HIERARCHY OF PROGRAM EVALUATION APPLIED TO VOLUNTEER LEADERSHIP DEVELOPMENT*

by

Claude F. Bennett
Specialist, Educational
Methodology and Evaluation

RESULTS ACHIEVED

"PRACTICE" CHANGE

"K A S" CHANGE

"REACTIONS"

RECIPIENTS INVOLVED

ACTIVITIES PERFORMED

INPUTS MADE

* Revision of presentation to a 4-H Youth Development Workshop on "Utilization of Current Knowledge About Voluntary Participation, Evaluation and Application to Problem Priorities," National 4-H Center, Chevy Chase, Maryland, February 14-18, 1972.

A HIERARCHY OF PROGRAM EVALUATION
APPLIED TO VOLUNTEER LEADERSHIP DEVELOPMENT

For Extension program personnel, "evaluation" is usually the last thing that comes to mind: last because most abstract formulations of the program development process place "evaluation" as the final step in a multi-phase sequence. Last in another sense, too; evaluation is often neglected in allocation of time and funds, in favor of "getting the job done." Is it possible that the frequent neglect of evaluation may be due partly to the conceptual approach mentioned, which places evaluation as the last phase?

If evaluation is usually an afterthought, planners of this workshop are to be commended in providing a time for state teams to make plans and assignments for an actual evaluation "back home." I am convinced that evaluation will never take its rightful place in education, including Extension, until plans for evaluation are specifically laid prior to implementation of the educational program.

KINDS OF EVALUATION

Although we are just beginning the "evaluation section of the workshop," it cannot be denied that a major part of our formal activity so far has been purposely evaluative.

In the broad sense of the term, "evaluation" is assigning a degree of value to something. Identifying a problem concerning volunteer participation, then, is a type of evaluation; some situation is defined as bad, or in need of improvement.

In the same way, state team have selected one or more possible or probable solutions for each problem. This, too, is a type of evaluation, as some potential solutions have been tentatively assigned more value than others.

And now the focus of the workshop turns to selecting an evaluation task for planning here, and for carrying out at home. What is suggested now is for state teams to plan for a more systematic evaluation--more toward the "scientific research" end of the continuum than you have done so far.

Perhaps a framework for selection of an evaluation task would assist our thinking at this point. Since any set of "pigeonholes" is somewhat arbitrary, this framework may not be the best, but it will serve to focus group discussion, and hopefully improve communications.

We can raise at least three kinds of evaluative questions in Extension:

First, "situational evaluation." This type of evaluation basically defines problems.

Second, "process evaluation." This is an evaluation of program planning, recruitment, administration and utilization of knowledge gained from educational programming.

Third, "program evaluation." For example, we could ask how much a 4-H volunteer program does to improve a situation. This question implies at least a bench mark or longitudinal study.

Thus, you may wish to plan for collection of further data to check whether you have correctly defined the problems you have already identified as state teams. Or, you could plan an evaluation of a new extension process, such as how roles are allocated to personnel. Or, thirdly, you may wish to plan for collection of data to assess the results of a new program.

Frutchey (1957) has defined evaluation as (1) collecting information, as a basis for making a decision or judgment; (2) applying certain standards or criteria to this collected information, and (3) forming the judgment or making the decision. Using Frutchey's definition of evaluation, the task you have at hand includes:

1. Deciding what information to collect about a situation, a process, or a program's results.

2. Deciding to what criteria you will relate this information; i.e., ascertaining the general and specific objectives.

3. Collecting the needed information and applying it to said criteria or objectives.

A HIERARCHY OF PROGRAM EVALUATION

Frutchey's is one of the most concise definitions of evaluation I have come across. But being very abstract, it does not tell us what types of information to consider. I would now like to present another facet of the framework for selection of a specific program evaluation task. This is a hierarchy of types of program evaluation. The hierarchy serves to help select exactly what type of evaluation is intended (c.f., Catalanello and Kirkpatrick, 1968).

Let me try to explain the hierarchy in general terms (see page 12). Type I in the hierarchy at the left indicates the easiest type of evaluation but unfortunately, also the least useful; the hierarchy proceeds to the right, through more and more difficult and complex types of evaluation, which are at the same time more and more useful. Type VII in the hierarchy is a master plan for evaluation. Administrators, program leaders, and evaluators can pick and choose from among these types of evaluation, basing their choice upon the size of the project envisioned, the resources allocated for evaluation,

and competency of the evaluation staff. Certainly this typology is an over-simplification, but the advantage of a model is to simplify reality so that we can get some kind of a "mind-hold" on it.

Six levels of evaluation are added in cumulative fashion as we go from left to right. Notice how a means-end hierarchy (Suchman, 1967: 52-59), is built into this hierarchy of evaluation. If we are to accomplish anything, we must expend some effort (input). Not any kind of effort, but educational activities, and we must reach an intended audience (recipients). But if the recipients don't appreciate the activities (reactions) they won't be motivated to learn anything ("K A S" change). If their learning is toward educational objectives, do they apply this learning ("practice" change)? And are they really better off (results achieved) if they follow practices we recommended? Thus we climb the hierarchy in order to reach the goals of the Cooperative Extension Service.

Let's take a detailed look at each type of evaluation beginning with the simplest type of evaluation, Type I, or "inputs made." This is a "reporting type" of evaluation which indicates how much time and money is being spent in the program. A Type I evaluation may also indicate how many staff are employed, their educational characteristics, etc. Is this really evaluation? Mostly it is just description, but it is also evaluation if we assume at least some correlation between how much effort is spent and how much is accomplished. (We are certainly willing to assume this correlation in adding personnel to a program.) But more important, we also need "inputs made" data for cost-benefit analysis within the "higher types" of evaluation to be discussed below.

A somewhat better evaluation is Type II, which besides telling us how much effort is expended also describes by means of written and audio-visual accounts the kinds and frequencies of activities performed in the project. This type of evaluation generally assumes that the activities held lead to a positive result. For example, who would deny that taking inner-city children on a field trip to a farm is beneficial? The trouble with this type of assumption is that it may not be true in certain situations, and moreover the question remains as to the extent of the benefits obtained. Of course, one variable in the extent of benefits is the number of recipients of the program and their degree of involvement. Therefore, we generally employ a Type III evaluation to gauge not only the number of those clientele or recipients who are directly contacted, but also those who are reached indirectly by mass media and by message diffusion. Again, it is common to assume that the more recipients, the better the program, perhaps even qualitatively. Of course, it is important that the audience have the "right" social characteristics, e.g., disadvantaged, urban, etc., in terms of the objectives of the program.

So far, we have discussed types of data which are considered as objectively measurable; i.e., they are relatively easy to count. These data may not by themselves tell us a great deal, but they are comparatively easy to gather.

Often, we wish to add at least a fourth level of evaluation, "reactions," in order to seek confirmation of the hunch that given activities are helpful. A Type IV evaluation adds to the previously discussed types of data "reactions" type data, or what people say about the activities. Such data should be representative of the audience, and also be quantitative, so that intensity of "pro" or "con" reaction is measured. The audience, the staff of the project, or

panels of observers are asked: "What is your attitude toward the project?" "Do you approve of it or disapprove of it? Why?" Valuable as such data may be, it may not indicate progress toward educational objectives. So, we may wish to press on to evaluation Type V, "change in knowledge, attitudes and skills."

Type V evaluations provide data to ascertain whether the activities are simply entertaining, or whether they really bring about the development of the group or individual, in terms of the educational objectives of the project. Type V evaluations show with quantitative data (a) whether changes have occurred in specific attitudes, knowledge and skills, (b) the direction of such change, and (c) the distance or extent of change. Data is usually collected within teaching-learning, or test situations. But the knowledge, skills and attitudes to be acquired are usually considered in "practical" education as just steppingstones to changing a behavior pattern.

Concern with quantitative data on acquisition of desirable patterns of behavior brings us to Type VI, the "practice change" or adoption type of evaluation. Continuous practice of the 4-4-3-2 nutrition rule, or of certain safety measures are but two examples. Of general interest in adoption studies is whether the behavioral change is really innovation, or just a modification of what is common practice. Duration of any change made is also of general interest.

Finally, we reach the "results achieved" or the Type VII evaluation, including "ultimate aims." These "ultimate aims" generally include concepts of personal and social development such as self-confidence, concern for the group, etc. This highest type of evaluation, VII, permits a cost-benefit ratio to be estimated, through comparing quantitative data on "results achieved," with "inputs made."

Although the lower types of evaluation may not have good measures on benefits received through the program, and thus have no way to compare such benefits with the cost of the project, every evaluation eventually has to answer the question, "Considering the total resources required, is the present program recommended for further work toward the solution of the problem selected for attack? If not, what approach is suggested?" So whether or not there is valid, quantitative data to support a cost-benefit ratio analysis, this final judgment regarding a program must be made.

DESCENDING THE HIERARCHY

We go up the "steps" to reach Extension goals, but we must come down the steps in order to evaluate Extension programs. That is, effective evaluation depends upon starting with the top steps and their interrelationships, and logically working our way down. Why? Because the principle to observe is this (Suchman, 1967): you can't evaluate change at "steps" I through VI without knowing or assuming the relationship between change at the "step" in question, and all the "steps" above it in the hierarchy!

Climbing the steps to reach Extension's goals is pretty clear. But the necessity of descending the steps in evaluation may require examples, which are provided in an application of the hierarchy in the next section of this paper. To demonstrate the necesity of descending the steps in evaluation, I shall try instead to climb the steps.

EVALUATION TYPES APPLIED TO LEADER ORIENTATION

To exemplify the use of the hierarchy, I would like to use Dolan's (1969) phases of leadership development as a point of departure* Let's take Dolan's

* The phases are: leader identification, leader selection, leader orientation, leader training, leader utilization, leader recognition, leader evaluation.

phase three, "Orientation of Leaders," and see how this hierarchy can be used in selecting an evaluation task. In the questions raised below, Extension orientation is evaluated, rather than the volunteer leaders. Also, only the added higher "level" of evaluation of each successive type of evaluation is discussed.

Type I. How much time and money do we expend in orientation of new leaders, and is this the right amount of resources (time and money) to expend? We can't answer whether the amount of input is correct without going to a higher type of evaluation, ** unless we know or are willing to assume how effective our orientation activities are.

Type II. In what kinds of orientation activities do we engage new leaders, and are we having the correct number of these activities? Again, we can't answer this latter question without assuming or knowing how effective the orientation activities are.**

Type III. With how many leaders do we allocate certain amounts of orientation time and activity? What are their characteristics, and are the proper number of such leaders receiving our orientation activities? Once again, an answer to the last question rests on the known or assumed quality of the orientation activities.**

Type IV. How do volunteer leaders react to their orientation activities? Do they think they get enough orientation? Too much? Notice that the answer to this question helps

** Of course, there are additional criteria needed to fully answer this question, such as measures of efficiency and degree to which performance is adequate to the total amount of need (Suchman, 1967: 63).

answer the three preceding questions. For example, suppose we accept reaction data as a valid evaluation. If it is the case that volunteers say we are giving too much orientation, then the more such activities a leader is exposed to, the more leaders exposed, and the more time spent in providing orientation, the less effective our program is!

Type V. What developments in the volunteers' knowledge, attitudes and skills are a consequence of our orientation activities? Again as with "reactions" data, "K A S change" data allows us to evaluate our "inputs made," "activities held," and "contacts made." But notice that we need "practice change" and "results achieved" data in order to evaluate "K A S change." We can measure whether "K A S change" occurs without knowing whether the skills, knowledge and attitudes learned are really the proper skills, knowledge, and attitudes.

Type VI. Do the attitudes, skills and knowledge gained through orientation help the volunteer to assist and interact in recommended ways with 4-H participants? As stated before, only when we have an answer to this question can we begin to evaluate changes in the volunteers' knowledge, skills and attitudes, brought about by the orientation program.

Type VII. Do the prescribed volunteer roles help achieve ultimate aims with 4-H youth, e.g., a better self-concept, higher citizenship aspirations, etc.? Again, it is only as we answer this "ultimate" question that we can evaluate the changes in patterns of behavior (practices) brought about by the orientation program (via changed skills, attitudes and knowledge).

CONCLUSIONS AND IMPLICATIONS FOR A STRATEGY OF EVALUATION

1. Any evaluative question within the hierarchy is related to, and has implications for, all the rest of the questions.
2. Evaluation at the lowest levels of the hierarchy is rather meaningless unless we either know, or make assumptions about the answers to "higher" evaluative questions.

   For example, it is meaningless to ask "Are we reaching enough volunteer leaders with our orientation program?" if we have no idea of:

   - How positively new volunteers react to the program.
   - Whether knowledges, attitudes and skills gained through orientation help the volunteers perform their expected role.

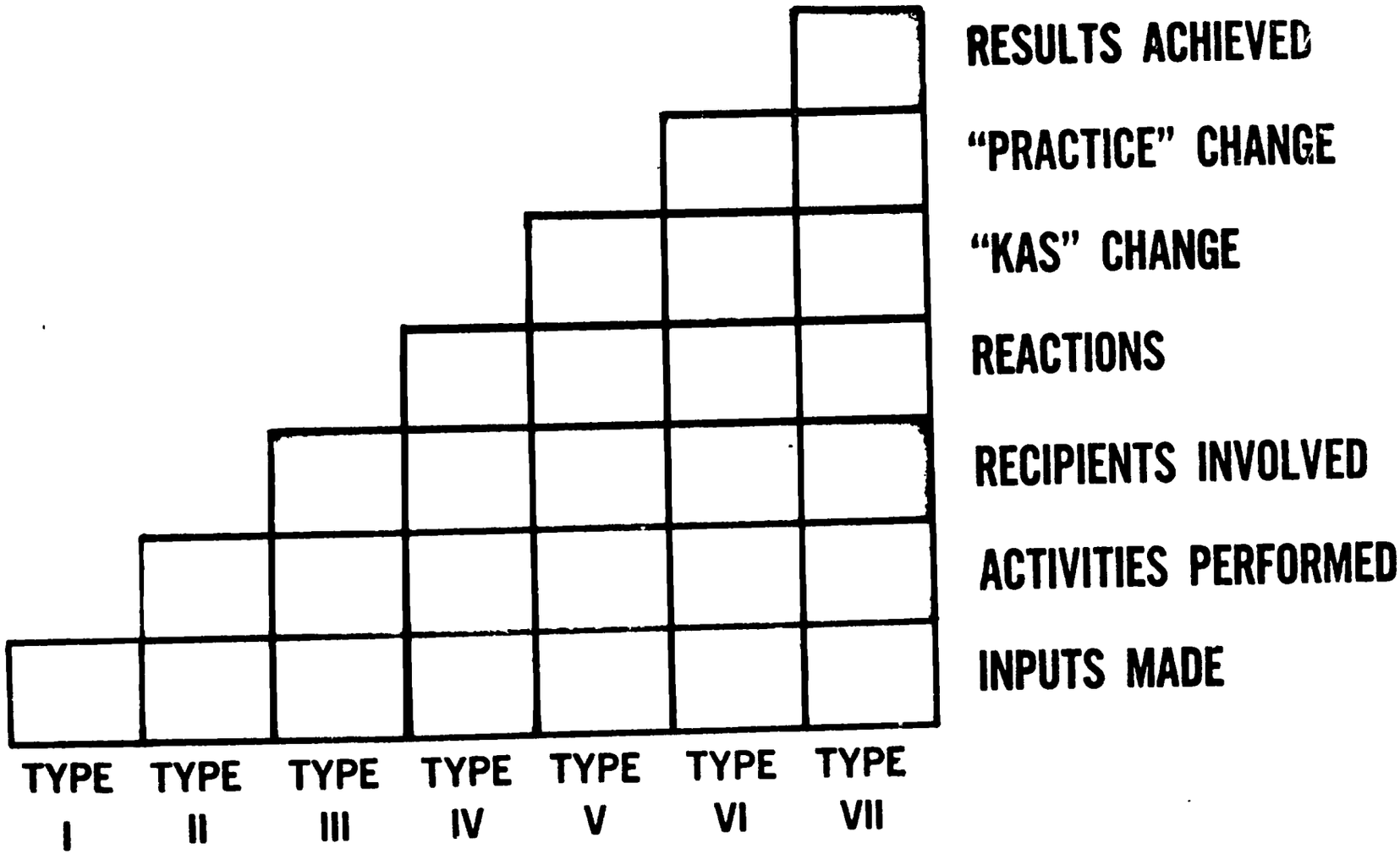
3. If a higher level evaluative question is answered positively through research (e.g., the knowledge and skill gained in orientation actually do help the volunteer assist in recommended ways), then the achievement of lower-level objectives can be interpreted as progress (although the progress may not be sufficient to meet the entire need).

   For example, we could conclude that if a greater proportion of our volunteer leaders are exposed to our orientation program, the more our

leaders will, in fact, assist in recommended ways.

4. Therefore, Extension should stress evaluative studies at the higher levels. These evaluation studies can help provide a basis for interpretation of less costly, lower types of evaluation.

A HIERARCHY OF PROGRAM EVALUATION



REFERENCES

Boyce, V. Milton

1971 A Systematic Approach to Leadership Development. Washington, D. C.: Extension Service, USDA

Catalanello, R. F. and Kirkpatrick, D. L.

1968 "Evaluating training programs - the state of the art," Training and Development Journal (May): 2-9.

Dolan, Robert J.

1969 The Leadership Development Process in Complex Organization.
Raleigh: North Carolina State University

Frutchey, Fred

1957 "Evaluation - what it is."
Pp. 1-5 in Byrn, Darcie (Ed.)
Evaluation in Extension.
Topeka, Kansas: H. M. Ives & Sons, Inc.

Suchman, Edward A.

1967 Evaluative Research.
New York: Russell Sage Foundation