

# innovative projects evaluation

---

Robert Brack

Considerable pressure is being exerted on adult education agencies to meet increasing and varied demands. New client groups, an increasing interest in continuing education, the knowledge explosion, and new educational technology are only some of the factors creating the ferment in adult education. No longer can we rely on the tried and true approaches. Greater experimentation will be the rule rather than the exception. How, then, can we know which of these new approaches works?

One way to tell, of course, is to rely on intuition. Some experienced adult educators may be able to select and sort the variety of approaches in programming without systematic evaluation. Most of us will have to rely on a more formal appraisal to make our decisions.

We're concerned here with the special kinds of considerations that evaluators must be aware of in assessing the value of innovative projects. There are, of course, considerations associated with any evaluation of ongoing projects to be taken into account as well.

There are kinds and degrees of innovation. One could argue that no project is ever completely replicated and conversely no project is ever entirely innovative. We are always making some adjustments in our projects, even when we consider them to be performing well. A project would be considered innovative if:

1. It's being offered to an entirely new audience (art classes for inmates in a penitentiary).
2. A different methodology is being employed (computer-assisted instruction for farmers).
3. The context in which the project operates has a new dimension (introduction of paraprofessionals into roles previously performed by professionals).

Trow says,

An innovation is a break with routine and habit; it disrupts unreflective ways of thinking, feeling, and behaving; it re-

---

*Robert Brack: Associate Professor of Continuing Education and Assistant Director, Extension Division, University of Saskatchewan.*

---

quires a heightened measure of attention and interest in the matters at hand; it forces the participants, and especially the creators, to think in fresh ways about familiar subjects, to reconsider old assumptions.<sup>1</sup>

## **Kinds of Considerations**

Evaluating innovative projects is different from the evaluation of ongoing and more traditional projects. There is usually heightened interest in the results of the study, the politics surrounding the project are more dynamic, and in many cases the staff are new on the job. While some of the following considerations also apply to studies of ongoing programs, they have special import when evaluating innovative projects.

## **Specifications of Outcomes**

Many creative ideas in programming get moved to the delivery stage before anyone has specified what the project's supposed to accomplish. While the developer of the project may have some notions of what kinds of outcomes he's expecting, these often aren't written down. The project may also be at a stage where it's not well defined and is changing as it unfolds. Conceptualizing a project and its counterpart in reality may be two different things. The evaluation design must be able to accommodate this dynamic state of development. Charters and Jones suggest that elaborately designed evaluation studies may sometimes end up appraising non-events where the "experimental" and "control" situations don't describe how they differ or even certify that they do.<sup>2</sup>

Pincus says, ". . . it is impossible to judge the merits of an innovation unless we have substantial information about how, and even if, it was implemented."<sup>3</sup>

Cain and Hollister feel that a major problem in designing evaluations of innovative programs is one of holding to an initial design concept long enough for an evaluation to be completed. If the design of the evaluation provides for a wide range of coping with variability in treatment variables, they feel the evaluation is likely to be more relevant.<sup>4</sup>

## **Levels of Performance**

Innovative projects have no forerunner. No track record exists for us to draw on in deciding levels of performance to expect the project to accomplish. What, then, can the evaluator do in setting performance levels?

If a prototype of the project has been developed, as in the case of learning materials, some data will be available from the testing phase. Some projects may have small-scale trial runs before the full-blown project is launched and these can provide some hunches about performance. We can examine projects that tend to be similar to the innovation to help in determining performance levels. We can have specialists, administrators, and sometimes client groups help in determin-

ing at what level they expect the project to perform. Of course, there's always the option of not specifying what level of performance is expected until the data are in. This has its pitfalls. A tendency exists to want to set the expected levels at the actual levels.

When there has been a thorough research and development phase before installation, the relationship between input and output may be fairly accurately guessed. But it's been suggested that research and development and mass operation phases tend to go on simultaneously in social programs and the needed research on the relationship between input and output hasn't been done yet.<sup>5</sup>

#### *Side Effects*

All programs produce effects in addition to those stated as goals of the program. But with established programs, we have pretty well sorted these out and incorporated positive side effects as legitimate goals for the program. The negative side effects are presumed to be minimal and can be ignored.

But, what about brand new programs? Should the evaluator concern himself only with stated outcomes for the program or be on the lookout for additional effects? Scriven suggests evaluators collect evidence about as many outcomes of a program as is feasible in the study and not just outcomes identified by the designer of the program.<sup>6</sup>

#### *"Hawthorne Effect"*

The tendency for new and experimental programs to succeed may well be attributed to the "Hawthorne effect." The name is derived from the famous experiment at the Hawthorne plant of Western Electric Company that showed increased productivity was due to the experimental environment itself rather than any of the independent variables used in the experiment.

One can easily see how this effect can influence the success of programs. In fact, Trow suggests that innovations have intrinsic qualities almost without regard for their outcomes. The reduction of boredom and the change in routine and habit can be enough to free our minds and energies for more productive work.<sup>7</sup>

In addition to the extra effort that programmers may contribute to the innovative project, the same kind of effect can apply to the participants in the project. Those in experimental projects often "try harder" so the project won't fail. Sometimes we may also find administrators lending more support to the innovative projects than would normally accrue to the ongoing program. Consequently, the program's effects are improved beyond what would occur under normal conditions.

#### *Sleeper Effects*

Evaluation studies are normally conducted within a time constraint. The evaluator needs to deal with the question of

whether the results of the project have manifested themselves in the data collected.

Some programmers will argue that the results of the project just haven't shown up yet—the sleeper effect. This argument has considerable validity when projects are aimed at changing value systems and customs. It should be clear to the evaluator what evidence can be expected within a specified time. If the effects of the project can't be expected within the time constraints of the current study, then an adequate evaluation can't be conducted. Perhaps a follow-up study at a later date has to be done.

### *Evaluator— Programmer Relationship*

Although interaction takes place between the evaluator and the people involved in conducting the project, with innovative projects this relationship will have different dimensions. The developers of the innovation may be more defensive; more may be at stake for the staff involved. On the other hand, the opportunities for the evaluator to help in improving and reshaping the innovation may be welcomed.

Early in the development of a design for evaluation, an understanding of the roles of the evaluator in relation to the programmer must be established. The programmer should be able to recognize how the evaluator will help rather than hinder him. Mutual confidence provides a good situation for conducting the evaluation.

Where the programmer views the evaluator as someone interfering with his territory, it will be more difficult to conduct the evaluation and get data sources. Imperative procedures such as getting participant lists may become difficult. Caro says that implicit in the evaluation role are attempts to discover inefficiency and to encourage change and there would be a tendency to conceal inefficiency and resist change. He further suggests that the evaluator establish effective ties with those who make key programming decisions.

One of the decisions to be made by the administrator is whether the evaluator should be an "insider" or "outsider." Tripodi and others identify the advantages and disadvantages of each arrangement.<sup>9</sup> Principally, the "insider" would have less difficulty with entry problems, whereas it would be expected that the "outsider" would be more objective.

### **Some Design Strategies**

As in all evaluation studies, the purposes for doing the evaluation of the innovation must be clearly stated. The users of the evaluation data need to be identified so that the study can be conducted in such a way that it's timely, relevant, and valid for their purposes. Probably the most helpful differentiation is to determine whether the evaluation is formative or summative.

The distinction between these two kinds of evaluation is that formative evaluation has the role of discovering deficiencies

and weaknesses while the project is under development. Summative evaluation is concerned with making judgments about projects after development is complete.<sup>10</sup> While such a distinction is useful in helping to design evaluation, most programs, don't seem to function this way. Evaluation is used at many points along the way, and may often serve simultaneously in a formative and a summative way.<sup>11</sup>

Another dimension that Gage feels is necessary is between the evaluation of an innovation in its general form and the evaluation of how it's working in a particular and local situation. In the first case, one is concerned with a design that will be able to say something about the performance in a general way. In the second case, the design has to determine how well it's performing in a particular system.<sup>13</sup>

---

**Evaluating innovative projects is different from the evaluation of ongoing and more traditional projects. There is usually heightened interest in the results of the study, the politics surrounding the project are more dynamic, and in many cases the staff are new on the job.**

---

Steele says that evaluation strategies are needed that work well in evaluating programs with high visibility, are new, and are controversial—particularly programs with a broad base of involvement of governmental agencies and the public.<sup>14</sup>

Keeping in mind the special circumstances surrounding innovations and the various roles that evaluation may perform, the evaluation designer must select an approach appropriate for that specific situation. In the current literature on evaluation, there are many models and approaches to evaluation. The following four have been selected because of their relevance to the design of evaluation of innovations.

*Controlled  
Field  
Experiment*

Many researchers consider a controlled research design with sampling and randomization as the only method of evaluating an innovation. Equally as many practitioners consider such a design much too rigid and impractical for field conditions. The fact is that some situations may lend themselves to a design that tries to have controlled, comparative experimentation. Stanley says that educational experimentation can be as ongoing, flexible, and sequential as the cleverness of the evaluators allow it to be.<sup>15</sup> In fact, in the monograph prepared by Campbell and Stanley, 16 experimental designs for research are described illustrating the flexibility of the experimental and quasi-experimental designs for evaluation studies.<sup>16</sup>

*Goal Free  
Evaluation*

Scriven has provided evaluators with some questions about doing evaluations based on the stated goals of a project.<sup>17</sup> His original concern was with side effects and their separation

from goals. He argues that the purpose of the evaluation should be to determine the effects of the project, whether they're intended or not. In carrying out goal free evaluation, he feels the less the external evaluator hears about the goals of the project, the less tunnel-vision he'll develop, and the more attention he'll pay to looking for actual effects.

Goal free evaluation may have particular relevance for doing evaluations of innovations. The stated goals of innovative projects are often, at best, stated in such grandiose terms, or so vaguely stated, that the evaluator could infer almost any outcome as a goal of the project. The search for data on side effects, both positive and negative, would likely be more thorough. Also, goal free evaluation doesn't have to take into account changes and shifts in goals as the project unfolds.

But, there are some questions about the practicality of doing goal free evaluation. Are evaluators sufficiently trained to avoid simply imposing their own goals for the project in place of the programmer's? Another problem is providing a focus for the collection of data. One can't observe and collect data about everything. Some suggest that the goal free evaluator may be more threatening to the producer of the project. There's also the possibility that programmers may well use goal free evaluation as an excuse for not specifying goals for their project at all.

### *Transactional Evaluation*

Rippey describes an evaluation strategy he calls transactional evaluation.<sup>18</sup> It involves not only the protagonists and the designers of the innovation, but also a representative sample of people likely to be affected adversely or disturbed by the consequence of the change. The model is built on resistance to change and its consequent threat to roles. Rippey suggests that to employ such a strategy, the evaluator must be familiar with, and either skilled at or assisted by people skilled at, interpersonal relations.

The involvement of both protagonists and antagonists of the change and the continual evaluation of both the anticipated and unanticipated consequences of change are the strengths of the model. He summarizes the contributions of transactional evaluation as:

1. Formative evaluation design is improved through the involvement of a wider range of opinions and values in the evaluation design.
2. Increased organizational efficiency and greater program benefits result because of attentiveness to potential role threats.
3. The concern of the evaluator for human values as well as program outcomes places him in a better relationship with personnel involved in the change, bringing greater honesty of interchange and thus more valid data.

4. Involvement of a wider range of interested personnel in evaluation leaves a residue of organizational and evaluative skills that are a potential for the organization, persisting beyond either the termination or the solidification of the original change.

**Adversary Model**

The adversary model as described by both Levine and Kourilsky is somewhat like Rippey's transactional model.<sup>19</sup> However, the Levine approach and the Kourilsky approach are different. Levine tends to use the methodology of legal proceedings in his model. Basically, he suggests, an adversary who should provide the "cross-examining" function and who would have access to the same sources of evidence as the primary investigator. He proposes that some group compile and classify the various types of evidence.

The Kourilsky model uses two evaluators—one who's an affirmative evaluator and another a negative evaluator. Each prepares cases for and against a proposal. The information from both aspects is provided to the decision maker. Considerable interaction among the two evaluators and the decision maker is suggested before a decision is made. She suggests that such an approach to evaluation is especially useful when a policy decision involving large amounts of resources is at issue.

**Summary of Models**

The following table identifies some of the unique characteristics of these four evaluation strategies. Its purpose is to contrast the four approaches. For those wishing to explore in more detail any of these approaches, the articles identified in the bibliography should be examined.

Table 1. Comparison of four evaluation models.

|                         | Controlled field experiment   | Goal free evaluation  | Transactional evaluation  | Adversary model  |
|-------------------------|---|---|---|--|
| Key emphasis            | Variables manipulated and their effects on other variables observed | Both intended and unintended effects of the innovation observed           | Includes people affected by the innovation as well as the designers | Both sides (protagonists and antagonists) of the innovation included     |
| Special skills required | Knowledge of research design required                               | Wide array of skills required   | Human relations skills required                                     | Skills in presenting case to decision makers required                    |
| Major advantage         | Conclusions drawn based on objective data                           | Allows for shifting of goals midway in a project and all effects examined | Acceptance of innovation by persons involved                        | Decision maker hears both sides and receives a wide array of information |

The literature on evaluation in the last few years is rich with descriptions of models and approaches to evaluation.<sup>20</sup> A variety of factors come into play as a design for evaluating an innovation is planned. Many of these are factors the evaluator has little or no control over:

1. The time constraint—an evaluation study is useful only when its results are available to decision makers when they need the information. Some designs take longer than others to complete.
2. The cost constraint—full-blown evaluation studies are costly. There's a cost effectiveness factor to consider for the evaluation study itself.
3. The experience and training of those charged with conducting the evaluation. It will be an agonizing experience for those involved if the design requires expertise that's not available to the evaluators.
4. The extent to which program staff are affected by the project or changes in the status of the project.
5. The context in which the program is operating. Projects with high profiles or controversial in nature need to be handled differently than those that are less contentious.
6. The extent to which the objectives of the project are specifically stated. An evaluation design for projects with well-defined objectives will be different than one where objectives are more generally stated or evolving as the project unfolds.

### Selecting Evaluation Model

Evaluating innovative projects can be a challenging, exciting undertaking. But, adult educators must experiment and be innovative in meeting the needs of evaluation. Some new approaches described here must be tested in the adult education field.

It's unlikely that one design or one approach will be appropriate in meeting the needs of all evaluation studies. The versatility of various models of evaluation is realistic if one chooses to be eclectic. The selection of elements from a number of models will probably provide the most practical approach.

### Footnotes

1. Martin Trow, "Methodological Problems in the Evaluation of Innovation," in *The Evaluation of Instruction*, M. C. Wittrock and David E. Wiley, ed. (New York: Holt, Rinehart and Winston, Inc., 1970), p. 291.
2. W. W. Charters, Jr., and John E. Jones, "On the Risk of Appraising Non-Events in Program Evaluation," *Educational Researcher*, II (November, 1973), 5.
3. John Pincus, "Incentives for Innovation in the Public Schools," *Review of Educational Research*, XLIV (No. 1, 1974), 134.
4. G. G. Cain and R. G. Hollister, "The Methodology of Evaluating Social Action Programs" in *Evaluating Social Programs*, Peter H.

- Rossi and Walter Williams (New York: Seminar Press, 1972), p. 131.
5. Joseph S. Wholey and others, *Federal Evaluation Policy* (Washington, D.C.: The Urban Institute, 1970), p. 96.
  6. Michael Scriven, "Prose and Cons About Goal-Free Evaluation," *Evaluation Comment*, III (December, 1972), 1.
  7. Martin Trow, "The Evaluation of Innovation," in *Readings in Evaluation Research*, Francis G. Caro, ed. (New York: Russell Sage Foundation, 1971), p. 82.
  8. Francis G. Caro, "Issues in the Evaluation of Social Programs," *Review of Educational Research*, XLI (April, 1971), 87-114.
  9. Tony Tripodi, Phillip Fellin, and Irwin Epstein, *Social Program Evaluation* (Itasca, Ill.: F. E. Peacock Publishers, Inc., 1971), pp. 122-25.
  10. Michael Scriven, *The Methodology of Evaluation*, AERA Monograph Series on Curriculum Evaluation, No. 1 (Chicago: Rand McNally & Co., 1967).
  11. Janet P. Moursund, *Evaluation: An Introduction to Research Design* (Monterey, Calif.: Brooks/Cole Publishing Co., 1973), p. 9.
  12. N. L. Gage, *The Evaluation of Instruction*, M. C. Wittrock and David E. Wiley, eds. (New York: Holt, Rinehart and Winston, Inc., 1970), p. 169.
  13. Henry M. Bricknell, "Appraising the Effects of Innovations in Local Schools," in *Educational Evaluation: New Roles, New Means*, Ralph W. Tyler, ed. (Chicago: University of Chicago Press, 1967), p. 285.
  14. Sara M. Steele, *Contemporary Approaches to Program Evaluation: Implications for Evaluating Programs for Disadvantaged Adults* (Washington, D.C.: Capitol Publications, Inc., 1973), p. 3.
  15. Julian C. Stanley, "Controlled Field Experiments As a Model for Evaluation," in *Evaluating Social Programs*, Peter H. Rossi and Walter Williams, eds. (New York: Seminar Press, 1972), p. 70.
  16. Donald T. Campbell and Julian C. Stanley, *Experimental and Quasi-Experimental Designs for Research* (Chicago: Rand McNally & Co., 1966).
  17. Scriven, "Prose and Cons About Goal-Free Evaluation."
  18. Robert M. Rippey, "Can Evaluation Research Be Integral to Innovation?" in *Curriculum Evaluation: Potentiality and Reality*, Joel Weiss, ed. (Ontario, Canada: Ontario Institute for Studies in Education, 1972).
  19. Murray Levine, "Scientific Method and the Adversary Model: Some Preliminary Suggestions," *Evaluation Comment*, IV (June, 1973), 1-3.
  20. For a comprehensive review of additional evaluation models, see Blaine R. Worthen and James R. Sanders, *Educational Evaluation: Theory and Practice* (Worthington, Ohio: Charles A. Jones Publishing Co., 1973), pp. 209-17 and Steele, *Contemporary Approaches to Program Evaluation*.