

HOW TO AVOID USELESS PROGRAM EVALUATIONS

By Steven E. Miller*

Accountability may top everyone's worry list these days, but we all know that most program evaluations are simply a waste of everyone's time and money.

Program funders, for whom these reports are ostensibly written, seldom pay attention. Or at least seldom get enough value from the report to justify the amount of work, hours, and resources that went into it. Most significantly, neither are most of these reports very useful to the project staff. They almost never address the really important day-to-day operational issues that practitioners struggled through. The style of writing is usually formal and obscure. They may contain lots of data, but few insights. And, in any case, it is a final report – the information comes too late to make a difference.

The uselessness of program evaluation is most apparent if the project is a one-shot or short-term effort. If the project is one step in a series of building block steps, then a summative report on one might provide some guidance for those that follow. But even in this context, most project evaluation is not as valuable as it should be.

And the situation is likely to get worse. These days, government leaders are constantly harping on the need for “scientifically based research” (SBR) and “data driven (or evidence based) decision making.” The rhetoric is seeping down into state grants as well as foundation and corporate giving agreements. Schools are trying to figure out what they are supposed to do, but everyone feels like they're in the middle of a muddle.

Part of the problem is general confusion about four different contexts for data analysis: research, project evaluation, program feedback, and planning.

RESEARCH

Research is about proving the validity of a basic hypothesis. In education, a researcher's most important job is to test a theory about the relative degree to which specific actions, programs, or other factors lead to a specific outcome. For example, a researcher might try measure how much a particular style of pedagogy, or a particular type of block schedule, or a particular curriculum program contributes to increased student learning outcomes. Researchers investigate questions such as “what kind of professional development has the largest impact on improving student learning,” or “under what conditions does technology make the most cost-effective contribution to student learning.”

Research involves measuring the difference between a “subject” group and a “control” group. If done using a true “medical model,” people are randomly assigned to be either in the subject or control group, and no one – neither the subjects, controls, nor the researchers – know which people are in which group, in what is called a “double blind” situation. Research involves statistically significant numbers of participants repeating the activity enough times so that all possible confounding variables can be mathematically ruled out. In a school setting, this means that, theoretically, Superintendents, Principals,

teachers, students, and other involved people are randomly assigned to each other and to one or the other group. It means that the number of participants is very large, the intervention being tested is very specific, and the long-term impact is determined through repeated measurements for many years after the original experiment. In fact, in an environment as complex as a school a good research project is more likely to raise additional questions requiring additional investigation rather than provide definitive answers. Still, when NCLB supporters talk about the need to use “research-based” strategies, they mean that schools should use strategies whose effectiveness has been validated through this kind of experimentation and sophisticated analysis.

This kind of research is time consuming and labor intensive and requires a great deal of expertise. If done correctly it provides vital but incremental additions to our basic understanding of teaching and learning. There are, however, few K-12 school districts capable of conducting subject-control, double blind, large-scale research suitable for professional peer review and publication.

Of course, schools can collaborate with researchers, and are likely to learn a lot by doing so. But, unless a program is designed from the beginning as a research project, shoehorning a school project into a research model is more likely to lead to confusion than to clarity. The environmental controls and data collection required for “proof” are simply lacking. The school-based program may help students learn, but it won’t be able to convincingly validate the theory on which it is based – nor should it be expected to do so.

This is not to say that schools can’t do research-like activity. It is always a step forward when teachers use some version of “Action Research” to come up with and then test the impact of classroom activities. For example, the Math Coordinator in one school system wanted to show teachers that using a new middle school curriculum would improve student learning. She asked a representative group of teachers to use the new curriculum in some of the classes and to continue using the old approach in other classes of relatively equal ability. The positive difference in the new curriculum classes, measured by both test scores and class participation, was strong enough to convince everyone to switch. While these school-based “research projects” make great articles for educational practice magazines, they would be unlikely to survive the kind of scrutiny given to submissions to research journals.¹

PROJECT EVALUATION

What schools can, and must, do is project evaluation. This type of data analysis does not set out to prove a basic hypothesis. Local projects are about implementing basic theories, not validating them. Project evaluation takes the underlying theory as a given. What

¹ In fact, not only should teachers be released from the demand that they create full-blown research, they should also be excused from being close readers of the research literature. Typically, research journals are full of long, tedious material that – at its best – just adds a small increment to our knowledge base and raises many more questions than it answers. Instead, a better approach is for schools to provide teachers with one of the newsletters that regularly scan research literature and conferences presentations and then provide plain-language summaries from a practitioner’s perspective.

project evaluation is supposed to measure is how effectively the basic theory has been put into practice as shown by the degree to which progress is made towards the anticipated goals. At its best, project evaluation documents not only that the project's finances are in order and all the planned activities actually occurred but also that it had the impact that it was aiming for, meaning that when the project was over the desired behavioral or learning outcome changes had occurred.

It is true that a project needs to be able to explain why its action plan is a valid method of implementing the underlying theory. There does have to be a coherent and logical connection between the actions taken and the goal sought, sometimes called a "logic model" or "logic map." If the underlying theory or premise has been research-validated and if the proposed actions are a reasonable and logically defensible method of implementing that theory, then all a project evaluation can do is document how well the action plan was implemented. This is quite different from doing research about the "truth" of the theory.

This is not to slight the importance of project evaluation in identifying key aspects of best practice. It is the job of research to identify the factors that lead to increased student learning, such as high quality professional development, increased time-on-task, good nutrition, etc. It is the job of projects – and their evaluations – to discover, for example, the most cost-effective method of using technology to accomplish these factors.

The key point is that project evaluation is not formal research. But it is increasingly hard for educational institutions to keep them separate. The pressure for "scientifically valid" results is pushing grant donors and recipients to use outside evaluators. Unfortunately, many of the people in the "evaluation industry" – the people most likely to bid for and get hired to act as outside evaluators for school projects – are trained to be researchers. Their professional success depends on producing reports that look good to other researchers or that can be published in research journals. And funders want to have the strongest possible proof that the activities they funded directly led to the claimed results – a research level of analysis. A successful evaluator is someone who writes reports that funders find convincing. So even when school leaders are clear that they are not looking for a research paper, they are likely to end up with one.

Beyond a simple confirmation of what activities actually took place, project evaluation in education programs should focus on four things. (1) Were there measurable increases in teachers' knowledge and skill? (2) Were there changes in teachers' behavior – as seen through instructional practices, types of professional development attended, the nature of teacher room conversations, etc. (Sometimes, it is important to document changes in administrative practices and policies as well, since these set the context for teaching practice.) (3) Were there changes in student learning activities and other related behaviors? (4) Were there changes in student learning outcomes?

Exactly how each of these is measured, the "units" that are used to express the measurements, and the scale by which the changes are deemed significant will vary according to each project's unique situation. If possible, the measurement occurs during

at least three time points: pre, post, and “later” -- with “later” meaning after participants have had time to reflect and incorporate what they’ve learned into their practice.

PROGRAM FEEDBACK

On one side, project evaluation tends to get confused with research. On the other side, it tends to get confused with in-process feedback systems. One cause of this second type of confusion is that fact that both final/summative and in-process/formative data analysis are officially called project evaluation. But the two have different goals and audiences.

Project evaluation is designed from the “outside in” – starting with the financial auditing and other final report requirements. Program feedback is evaluation for the purpose of in-process decision-making. Feedback systems have to be designed from the “inside out” – starting with the needs of the project participants and leaders. In addition to tracking progress towards the stated goals, program feedback includes an analysis of what the participants have done or can do to increase the level of measured achievement.

While a good program feedback data system will usually also create the kind of data needed for a summative project evaluation, the reverse is seldom true. It is unlikely that a data collection system designed to serve project evaluation needs will also provide useful in-process feedback. If the priority is on producing a final report, project leaders need to explicitly create a separate system for on-going data collection and analysis – as well as structuring in opportunities for participants to reflect on the data and make course-corrections.

What kind of formative, “running record” data collection is needed? It is useful to provide feedback about individual progress as a way of deepening teacher understanding and documenting growth. But teachers almost always have a solid sense of a student’s general level of understanding and skills. Feedback is most useful if it raises the level of analysis to group trends and patterns and program-level developments. Are we doing what we said we would do on the planned schedule and with the anticipated level of participation? If not (which is almost always the case), what is taking longer than expected, and why? How does participant’s initial levels of knowledge and skill differ from what we expected? How much, of what, are participants learning? What impacts are beginning to appear? What have been participants’ feelings about our various activities, and why do they feel that way? If we are working with other groups, has the partnership been fully effective, and if not, why? Are we beginning to lay a foundation for post-grant sustainability by increasing internal staff capacity and leadership, by incorporating our project’s methods into standard operating procedures, by generating positive indicators of our impact, and by increasing community and decision-maker awareness and support of our efforts?

Being reflective and flexible is a difficult challenge. It takes work and time to regularly collect and analyze this kind of data. Most of us simply want to get to work and push forward, making intuitive adjustments as we go. It often necessary to make someone, a staff person or an outside consultant, explicitly responsible for program feedback. But formative data is only useful if it is used formatively. That requires building time and

resources into the initial project design and budget. That requires making it clear to everyone involved, even before you begin, that collecting and dealing with feedback is a basic part of the process. The good news is that most people are happy to participate if a useful process has been set up for it.

PLANNING

The fourth type of data analysis is the kind done for project planning. Sometimes called a “needs assessment,” the goal is not to prove a hypothesis or fine-tune an implementation effort. Evaluation of data for planning is shaped by the need to focus attention on a do-able component of an overall situation and to help shape – and then legitimize – the chosen program design. Planning data convincingly documents the existence of a problem, the relevance of a research-validated theory to its amelioration, and the coherence of the logic-map of actions that the project intends to follow. It is often also useful to document the likelihood that the people making the proposal can successfully carry out the action plan given local history and resources.

USING AN OUTSIDE EVALUATOR

Not every project needs to have an outside evaluator. But project leaders who are not fully confident about their ability to create effective data gathering and analysis systems would be foolhardy if they didn’t look for outside help, even if only on an occasional consulting basis.

At the same time, schools seldom use outside evaluators to maximum advantage. First, K-12 grant developers are often overly impressed by the need for “research validity” and the accountability requirements of the funders. As a result, they end up going along with evaluator’s inherent bias towards a research model. But they don’t have the resources to do a full-scale research project. So the program evaluation ends up feeling like an annoying irrelevance and the staff walk around making fun of – or being frustrated with – the extra and extraneous work it requires.

A second problem is that evaluators are usually brought in much too late. If they are not part of the project planning process from the very beginning, evaluation inevitably becomes an after-thought, something tacked on at the end. If brought in late, even the most well-meaning evaluator will have trouble not being seen as an outsider and a pain. Furthermore, if the evaluator is not brought in until the project is already up and running it becomes difficult to even collect the baseline data that will be needed for the final report, much less to be able to play a constructive formative role that wins the cooperation of the participants.

A good evaluator should be a fully participating member of the project design team from the beginning. Even before a program is designed or a proposal written, a good evaluator can help in several ways.

- * Help doing the needs analysis to develop a clear understanding of the problem;
- * Help identify research-validated insights about and strategies for dealing with the identified problem;

- * Help developing measurable goals;
- * Help design a data collect process that provides on-going feedback about key issues you want to track, as well as lays the foundation for required final reports;

When you are writing the proposal, a good evaluator can:

- * Help write or edit the goal statement and evaluation method so as to best meet the funder's requirements while preserving the core "inside out" data analysis design.

Once the project is funded and in operation, a good evaluator should:

- * Help periodically analyze in-process data, talk with project leaders and participants about its meaning, and do additional research if needed to explore ways to address any apparent shortcomings;

And as the project winds down, a good evaluator should:

- * Pull together all the data, write the final reports, and then meet with project leaders, participants, and other stakeholders to discuss the results.

In the largest sense, we avoid useless project evaluation by knowing what we are doing. And the needed knowledge includes an understanding of how to structure your project evaluation so as to get the most value for your time and money.

Good luck.

** Steven E. Miller is Executive Director of Mass Networks Education Partnership, a New England-based nonprofit education consulting firm that provides Data-to-Learning Programs, Technology Services, Curriculum Development Services, and Project Evaluation to K-12 districts.*