

8

This chapter explores how theory-based evaluation can be used to help conduct formative, reflective evaluations in educational settings.

Theory-Based Evaluation: Gaining a Shared Understanding Between School Staff and Evaluators

Tracy A. Huebner

Gaining cooperation from program staff is an ongoing problem for evaluators. This is true in schools and in a number of other settings. This chapter focuses on how theory-based evaluation can help address this problem in school evaluations, but its argument applies to a broad array of evaluations.

Five significant challenges plague school site evaluations. One of the barriers is staff's lack of receptivity to an outsider coming in to evaluate their programs. There are several reasons for this **lack** of receptivity. For example, teachers may have had negative experiences with evaluators in the past; evaluations may have led to the closing of a program. Or teachers may have felt that they were misrepresented in a final analysis. Whatever the reason, teachers unwilling to participate in an evaluation make the task difficult.

A second challenge for evaluators is that teachers are not trained in evaluation. Because they may not see or understand the relevance of this work to their own, they may be less than willing to participate in data collection or other aspects of an evaluation. For example, the teacher at Woodland who told an evaluator to do his job while she did hers did not understand her role in the evaluation. This is not unusual. Teachers often

Note: Earlier drafts of this chapter were presented to the Evaluation Task Force of the Harvard Children's Initiative and the American Evaluation Association Annual Conference in 1998. This chapter was generously sponsored by a grant from the Spencer Foundation. The author wishes to acknowledge the wonderful insights and **support** of Carol Weiss, Timothy Hacsí, Anthony Petrosino, Patricia Rogers, Barbara Neufeld, and George Madaus.

perceive the role of evaluation as outside of their arena; anything having to do with evaluation has nothing to do with them. Compounded with the challenges of teacher resistance and lack of training, evaluators are also burdened with the difficulties of evaluating complex school-based programs.

Third, schools have weak veins of communication. Oftentimes messages sent from the administration to the teachers about a new program or policy are interpreted in as many different ways as there are teachers. Because administrators spend so little time in classrooms, and teachers spend the majority of their time teaching independently, there is no method in place to clarify these different interpretations. Instead the separation between administrator and teacher exacerbates the communication problem.

Fourth, adding to the challenge of an already weakened system of communication is the fact that school-based programs are very complex. Many of the current school reform efforts combine both administrative and academic change (see, for example, the Accelerated Schools Project and the Coalition for Essential Schools).¹ Programs are asking teachers to change not only the way they teach but also the way they participate in the school community, for example, participating in school governance boards and interacting with the parents and business community. This means that the evaluator has to look at not only the academic development of students but also the culture and climate of the school. All components are intertwined, and disentangling them is a challenge. Communication and complexity are two issues that are difficult to handle but can be addressed by evaluators.

A fifth challenge beyond the boundary of control of evaluators is the fact that many programs are not in place long enough for us to really learn about their merits and weaknesses. All too often, programs are terminated after one or two years of operation based on the claim that student test scores did not increase. But research on school change tells us that it takes at least five years to observe substantive change (Fullan, 1991).

With all of these challenges in front of evaluators, how are we working to address them? Evaluators are turning to innovative approaches to involve educators more directly in the evaluation process. They are drawing on different models of evaluation to break down barriers between teachers and administrators, as well as between school staff and themselves.

Theory-based evaluation is one model that attempts to include program staff in the design and implementation of the evaluation. This model identifies the mechanisms, or the links, between the planned activities and the anticipated outcomes. This chapter explores theory-based evaluation as an approach to conducting formative, reflective evaluations in educational settings. Although the chapter looks specifically at school sites, the lessons learned can be applied to other programs as well. In practice, this model of evaluation helps schools clarify program goals, improve cooperation and buy-in in an evaluation, and encourage reflective practice.

Empirical Evidence to Support Theory-Based Evaluation with School-Based Programs

Information was gathered through interviews and the study of documents.² I examined four theory-based evaluations conducted by four different evaluation teams all adhering to two criteria: each program had an explicit model of its “theory” identified by the evaluator and program staff before the actual evaluation took place, and the evaluation tested at least one mechanism in addition to its ultimate outcome.³

I identified three major themes in their evaluations. First, evaluators reported that the process of using logic models as a way to identify the program theory helped both the evaluation team and the school staff clarify program goals. Second, by establishing a common understanding of the program’s goals and evaluating the program based on this shared understanding, school staff tended to be more cooperative throughout the course of the evaluation. Third, two evaluators in the sample believed that implementing a theory-based evaluation encouraged teachers to be more reflective in their own practice. (For more about the relationship between theory-based evaluation and reflective practice, see Huebner, 1998.) The matrix depicted in Table 8.1 identifies the evaluators and provides a brief description of each evaluation, the mechanisms tested, and the reported advantages to using the theory-based model. Following the matrix is a more detailed description of the programs and the themes identified in support of theory-based evaluation.

Helps Clarify Program Goals. The evaluators reported that theory-based evaluations not only helped establish a rapport between the evaluator and the school staff but also helped unify the staff. The initial process that helped pull all parties together was the activity of developing a logic model of the program’s theory. Evaluators reported that the act of designing logic models was perceived as an important step in building relationships with the school staff. Designing these models helped staff feel included in the evaluation, offered evaluators a chance to interact in a nonthreatening way with staff, and acted as an important tool, once they were developed, to help the evaluator and staff identify key questions for evaluation.

Together the staff and evaluator had a shared understanding of how the program was supposed to work at each level. Gaining clarity so that all parties involved in the evaluation were aware of the activities and the mechanisms leading to the intended outcomes was perceived as a critical first step toward a successful evaluation.

Owen and Lambert (1995) believed that their 1993 evaluation of a technology program that had been initiated in a fifth-grade curriculum was enhanced because it had begun with a logic model. The evaluators worked with the staff to define their theory of how the integration of computers would affect the existing curriculum. The model helped “identify the full range of program consequences and, *in addition*, plac[ed] the program within the context of the total school system” (p. 246).

Table 8.1. Matrix of Four Evaluations Using a Theory-Based Approach

<i>Evaluator</i>	<i>Description of Evaluation</i>	<i>Mechanisms Tested^a</i>	<i>Reported Advantages to Using a Theory-Based Model</i>
Owen and Lambert (1995)	Evaluate the introduction of laptop computers into a middle-school curriculum.	Links among (1) providing teachers with basic skills related to personal mastery of laptop computers, which (2) helps teachers develop a strong level of proficiency in programming skills, (3) provides more opportunities for students to learn independently at their own pace, and (4) integrates laptop computer use into the classroom.	Unintended outcome had broad implications for school goals—for example, implications for curriculum impact, professional development, and support.
Finnan and Davis (1995)	Evaluate the implementation of a whole school reform effort at the whole school and class room levels.	(Individual classroom models were tested only. ^b) Example, links among (1) planning and implementing grade-appropriate material, (2) providing one-on-one instruction for students in need, and (3) assigning at-home reading, which (4) results in increased student reading proficiency.	Teachers became more actively engaged in evaluation—collecting and analyzing data. Teachers reported greater awareness of how their classroom fit in with the overall philosophy and process of the school.

Beckford (1998)

Evaluate whether or not intelligence can be learned.

(At this point, only an impact evaluation is being done.)
Future links among (1) improving school leadership, which (2) leads to change in teacher attitudes and beliefs, which in turn (3) encourages teachers to use new assessment practices.

Staff grew to understand their program better.

Darling (1998)

Evaluate the integration of an eighth-grade core curriculum with state standards.

Links among (1) orienting students in reading, writing, communications, critical thinking, and independent learning, which (2) helps students build schema for understanding history.^c

Teachers were more actively engaged in evaluation design and implementation. They altered practices as a result of learning from the evaluation.

^a*Mechanisms tested* refers to the actual links in the program theory that were tested, therefore this column does not reflect entire program theories.

^bActual evaluation was not carried out because the school closed while the evaluation was under way.

^cDarling reports only the first phase of the evaluation at the time of this chapter.

Owen and Lambert's evaluation linked the technology program with the inner workings of the school as a whole. So instead of just focusing on how the technology program worked in isolation, the theory-based evaluation helped address questions that asked how the program was working *and* how it was working *in the context of* the school's overall organizational structure. For example, did the computer program support the math curriculum? Did it support the overall goals of the school to create an interactive learning environment for children? Owen and Lambert's theory-based evaluation clarified program goals and enabled the evaluators to see how they fit in with the overarching school goals—a very important component for evaluation of school-based programs.

This strategy was also used by Finnan and Davis (1995) in their evaluation of Berry Elementary School's approach to schoolwide reform. They reported that as a result of this approach teachers as well as administrators came together for the first time and worked toward a common understanding of what was happening in classrooms. Together the staff looked at how classroom activities related to what was happening schoolwide. Once this common understanding was established, the faculty and staff worked together with the help of the external evaluators to design an evaluation to meet classroom and schoolwide needs.

Darling and her evaluation team of teachers at Wenatchee Middle School used a theory-based model of evaluation to look at an integrated eighth-grade social studies and language arts curriculum (J. Darling, personal communication with author, Mar.–May 1998). They used the logic model as a way to determine the soundness of their program design as well as to select goals and activities to streamline the curriculum. The process of developing a logic model helped the teachers compare their student achievement goals with outcomes and modify the program according to their needs. In addition, the teachers wanted to use a model that was clear to all stakeholders in the school community. They believed that using a theory-based evaluation would enable them to provide a clear picture of their program and the evaluation to those not directly involved. They wanted this clarity so that other members of the school community would understand their work, understand their evaluation, and critique both.

Ian Beckford (personal communication with author, Mar.–Apr. 1998), like Darling, believed that the process of developing a model of the program's theory helped explain the program more clearly to both himself and the staff, and this clarity in turn helped ensure a more accurate evaluation. Although he talked positively about the process, Beckford was careful to warn of the challenges he had faced in using this model of evaluation. As a result of choosing to work with staff rather than independently, Beckford came up against many interpersonal conflicts regarding the evaluation. His struggle with school staff was not in creating an evaluation based on their program theory as much as it was in developing the model itself. This is not all that uncommon. Weiss (1997) notes that program staff often have difficulty articulating what their theories are.

Beckford's most current work, in 1998, with two educational programs, one a small-scale after school homework-tutoring program and the other a large-scale national school reform effort, involves identifying the program's theory with staff as a way to "work through with [staff] what they [think] the program is about" (Beckford, personal communication with author, 1998). As a result of this time spent "up front," Beckford felt that his clients and program staff with whom he worked were more attuned to the overall development and process of the evaluation.

Builds Cooperation and Buy-In, Which Helps Develop the Evaluation. Based on both Darling's and Beckford's observations of developing logic models of program theory, a good way to develop the models is with the assistance of program staff. The evaluator should take the lead role as facilitator. As a result of this common experience, the staff becomes more receptive and willing to participate in the evaluation. In the majority of evaluations reviewed, staff expressed an interest in raising the questions for the evaluation, collecting data for the evaluation, and analyzing the data.

The evaluations reviewed in this chapter include school staff as critical players in the development of these models. After all, who knows the program better than the ones participating in it day to day? In contrast, the evaluator is key in the development of the program theory because it is the evaluator who provides the viewpoint of the outsider—one who can see the forest for the trees (Patton, 1997). These blended perspectives help build a well-rounded picture of the program, its goals, and the mechanisms by which the program seeks to attain them.

Finnan and Davis (1995) wrote about using evaluation to link individual teachers' work with whole school goals and about the powerful connection it helped both teachers and administrators make. They wanted to engage teachers in understanding what was happening in their classrooms as well as in the school as a whole and how the two were interrelated. They also wanted to provide teachers the opportunity to take control and direct their own evaluation by examining the issues that they believed were most relevant to their teaching and their school.

Darling argued that data collection must be "practical, useful, meaningful and trustworthy" in order for members of the school community to find the evaluation credible. She believed that using a theory-based evaluation would help her develop a systematic strategy to meet this end. Her prior experience taught her that "no faculty seems to voluntarily and systematically evaluate their own program. . . . we have no practitioner-designed tools to fulfill our program evaluation responsibility; we have very few (if any) practical, useful, meaningful, trustworthy methods for systematically improving our teaching ourselves" (Darling, personal communication with author, Mar.–May 1998).

However, despite Darling's positive view of the merits of using a theory-based evaluation, the majority of her colleagues were less interested in the actual work of data collection and analysis and more in the development of

program models. Teachers were more interested in developing the model than in looking at its implementation. She wrote, “The front end [designing logic models] was really their only interest because basically everything else seems unimportant to them. [Teachers] want to develop and improve programs but are not interested in collecting or analyzing data in the traditional sense. . . . There was little interest in finding out about implementation, more interest was in development of a program, trying to conceptualize it in a practical way” (Darling, personal communication with author, Mar.–May 1998).

Darling’s impressions of teachers’ reactions to evaluation are grounded in only one experience. It is quite possible that the teachers were less inclined to participate in the evaluation for reasons beyond the scope of what this chapter covers. There is no direct evidence suggesting that the use of a theory-based model of evaluation inhibited or encouraged the teachers from Wenatchee to evaluate their curriculum any more or less than a different model.

Teachers do not often know how a school goes through reform and the role they play in the reform. The theory-based evaluations of Darling and Finnan and Davis helped stress these connections and make teachers aware of how the work they do in their classrooms relates to the overall goals of the department and the school. As a result, both Darling and Finnan and Davis believed that their evaluations received greater support from the teachers than they would have if the teachers had not been involved in the initial planning phase.

Encourages Reflective Practice. Another important benefit of a theory-based model is its potential to encourage *reflective practice* in educators. Reflective practice is the process by which teachers actively think about their work and how it affects their students. Based on this understanding, teachers can modify instruction to better meet the learning needs of the students. Teachers exist beyond the landscape of the classroom. A teacher is an individual who works with others to help deepen knowledge. This teaching role—working with others to increase knowledge and understanding—exists in many different sectors, for example, in the health profession when doctors work with residents to train them in their craft. This relationship also exists in community-based programs for youth and adults, where program staff work with clients to increase their knowledge and understanding on any number of issues. As a result of the widespread need for good instructional practice, evaluators must think about the benefit of encouraging reflectivity in general.

Dewey believed that teachers should engage in reflective practice in order to promote opportunities for student learning. He argued that reflective practice is intentional and is directed to the aim of making meaning from interactive teaching episodes. Why encourage reflective practice? Because the act of teaching, whether it is in a classroom or any other context, is a complex act. Reflecting on practice is one way to address myriad

intricacies faced by teachers as well as others engaged in working with individuals in a learning capacity. Teachers need to be able to sort through volumes of material in order to teach content to students. It is not enough for teachers to select materials at random for their subject. They must know how to cull and identify the most salient materials that are appropriate for their students and that accurately represent what they are trying to teach. The teachers who reflect thoughtfully about students and subject matter are the teachers who best design methods and materials.

Reflecting on data collected for an evaluation and modifying one's practice to address the needs of the learners is precisely what Schön had in mind when writing not only about reflective practice (1983) but also about his work with Argyris on theory-of-action evaluation (Argyris and Schön, 1974) ten years prior.

In Darling's evaluation, teachers worked together for four days, building their logic model for their eighth-grade social studies curriculum. Darling equated those four days as "four days of reflective practice and for most of us, exhausting and exhilarating." Darling thought that the teachers engaged more in reflection on action than reflection in action. Darling and her colleagues thought that reflection occurred outside the classroom. "When we are actually practicing, life is crazy. We have zero time to reflect. We're just scrambling—trying to get progress reports out, talking to parents, running off papers, tinkering with rooms, etc." (Darling, personal communication with author, Mar.–May 1998).

In Finnan and Davis's evaluation, teachers were encouraged to actively record and think about their practice and its relation to their overall goals both at the classroom level and the whole school level. Teachers engaged in thinking through their own theories of how their students would achieve their ultimate learning goals. Throughout the year, teachers were encouraged to reflect on how their students were progressing academically. Teachers recorded student performance, made observations regarding student work, and modified their practices to better address the needs of their students. In addition, teachers were encouraged to see the links between their own classroom interactions and the school as a whole.

Finnan and Davis's work suggests that involving teachers in an evaluation by encouraging them to look at their own classroom and to see how their work in the classroom affects the overall motion of the school toward reform provides a rich forum for reflection. The evaluators worked with Berry Elementary School to chart the change process at the school sites—the links between and within the classroom, small group, and whole school.

Not all teachers were able to be reflective on their practice and to modify it based on data collected in the evaluation. Some teachers fell into the act of engaging in a process-outcome evaluation. Instead of using student work to reflect on how their presumed mechanisms lead to improved student outcomes, they used the data and journal as a checklist to record

their activities. This is not surprising. Not all individuals are predisposed to being reflective. The development of a reflective practitioner is a combination of innate qualities and learned habits (La Boskey, 1994). Teachers predisposed to reflection are more likely to engage in the process of reflecting on their instructional practice and its effects on student learning. For others, reflection is something that they feel they do not have the time for or that it is an act in which they cannot engage without ample coaching.

Conclusions

This chapter presents data on only four evaluators engaged in theory-based evaluations in educational contexts. Although the initial findings of all four evaluators suggest the benefits of using this strategy, a more substantive investigation is required.

When developed and articulated by both the evaluator and the program staff, program theory provides an approach for mutual understanding of the intentions of the program, and it is a way of laying the groundwork for comprehending why an evaluation is useful and what kind of evaluation design is most beneficial.

Schools seeking to implement evaluations that are helpful to teachers in the classroom and schoolwide may choose to implement a theory-based evaluation. This model moves beyond the black box and provides information about how and why a program functions rather than judging whether or not the program works.

Evaluators using a theory-based evaluation with school-based programs believe that it helps clarify program goals and increase buy-in and participation in the actual evaluations. In addition, theory-based evaluation encourages teachers to be more reflective in their practice. This occurs through the process of collecting data and reflecting on the data, which is a way for teachers to evaluate their practice and their students' learning based on their instruction.

Notes

1. Information regarding the Accelerated Schools Project can be accessed on line at www.stanford.edu/group/asp/natlcenter.html. For information regarding the Coalition of Essential Schools, refer to www.ces.org.

2. I engaged in personal communication with three of the four evaluation teams, and in two instances I **spoke** with members who participated in the evaluations, such as teachers and a school principal. In all cases, I talked with the evaluators about the differences they perceived between using a theory-based model of evaluation and another approach. I also talked in depth with the evaluators, teachers, and a principal about the challenges and successes of using logic models as a way to interpret and evaluate a school-based program.

3. The criteria for the definition of theory-based evaluation were determined by Weiss, Hacsí, Huebner, Petrosino, and Rogers in 1998 at the Harvard Children's Initiative.

References

- Argyris, C., and Schon, D. A. *Theory in Practice: Increasing Professional Effectiveness*. San Francisco: Jossey-Bass, 1974.
- Finnan, C., and Davis, S. "Linking Project Evaluation and Goals-Based Teacher Evaluation: Evaluating the Accelerated Schools Project in South Carolina." Paper presented at the annual meeting for the American Educational Research Association, San Francisco, 1995.
- Fullan, M. *The New Meaning of Educational Change*. (2nd ed.) New York: Teachers College Press, 1991.
- Huebner, T. "Theory-Based Evaluation and Reflective Practice: What's the Link?" Unpublished paper, 1998.
- La Boskey, V. *Development of Reflective Practice*. New York: Teachers College Press, 1994.
- Owen, J., and Lambert, F. "Roles for Evaluation in Learning Organizations." *Evaluation* 1, 1995, 1, 237-250.
- Patton, M. *Utilization-Focused Evaluation*. Thousand Oaks, Calif.: Sage, 1997.
- Schon, D. A. *The Reflective Practitioner*. San Francisco: Jossey-Bass, 1983.
- Weiss, C. H. "Theory-Based Evaluation: Past, Present, and Future." In D. J. Rog and D. Fournier (eds.), *Progress and Future Directions in Evaluation: Perspectives on Theory, Practice, and Methods*. *New Directions for Evaluation*, no 76. San Francisco: Jossey-Bass, 1997.
- Weiss, C. H. *Evaluation: Methods for Studying Programs and Policies*. (2d ed.) Englewood Cliffs, N.J.: Prentice Hall, 1998.

TRACY A. HUEBNER is coordinator for comprehensive school reform at WestEd.