EVALUATION CAPACITY BUILDING IN A SCHOOL-UNIVERSITY PARTNERSHIP GRANT PROGRAM

After several years in which only superficial outcomes measures were used to evaluate grants funded by the Illinois Board of Higher Education (IBHE), the Illinois State University Center for the Study of Education Policy piloted a new approach by which real project assessment could be embedded in Improving Teacher Quality (ITQ) project designs when first funded in 2004. This approach applied logic modeling at both program and project levels. As a result, feedback about the program and the individual projects was integrated into evaluation at both of these programmatic levels. This article describes the details involved in seven years' work with the more sophisticated evaluation model. The ITQ program has moved from funding 26 Science, Technology, Engineering, and Mathematics projects with marginal evaluation plans to funding nine projects with solid, research-based logic models and aligned evaluation plans. The article also describes the difficulties involved in any attempt to compare one unique program with another, and instead uses the metaphor of the “investment portfolio” to both describe and assess the IBHE’s array of investments in the ITQ projects.

This paper outlines challenges faced by policymakers managing grant funds and describes a process whereby a state agency can build evaluation capacity to promote continuous improvement of project outcomes. In the current economic climate, discretionary grant funds are often a key resource to underwrite educational improvement efforts. Such funds are scarce and therefore precious. Both grantees and grantors are held accountable for producing measurable results from the investment of these funds. In this paper, we look specifically at Illinois Board of Higher Education (IBHE) grant funds for teacher professional development projects, distributed competitively under Title II of the federal Elementary and Secondary Education Act (ESEA). We make the case that building evaluation capacity and sound evaluation practice throughout the granting process can improve results and return on investment.

We first describe the ITQ program, the challenges of grant-making, and the potential utility of evaluation capacity building (ECB). We then propose a logic model for embedding evaluation throughout the granting process. Each component of the logic model is discussed in detail, including feedback loops for continuous improvement. Finally, we summarize the role a grantor can play in modeling good evaluation practice, thereby leading improvement efforts.
The Program Challenge

As in many states, public education represents the largest proportion of Illinois’ budget. In FY2011, education spending consumed 33.7% of general revenue funds (GRF) and 22% of all funds (including capital investment). For comparison, the next highest spending category was health and human services, consuming 23.1% of GRF. Most of this money is distributed to schools and universities through formula grants, in budgetary line items for various programs and operations. Very little of this funding is available for discretionary programs to foster creative projects in various locales.

An important exception is that a portion of federal funding that flows through state budgets is available as discretionary grants for special projects. One example is Title II funding within ESEA. Each year, a small percentage of this money (about 2% of Illinois’ allocation) is channeled through IBHE for competitive partnership grants between colleges and school districts in a program the state calls Improving Teacher Quality (ITQ). These grants are meant to underwrite creative professional development partnership projects that improve teaching and learning in classrooms around the state. In Illinois the bulk of the projects are devoted to projects in science, technology, engineering, and mathematics (STEM) disciplines.

Grants are awarded on a three-year cycle through a competitive Request for Proposal (RFP) process. The first year is an “open” competition in which new and existing partnerships submit proposals for three years of project funding. All proposals are reviewed using a scoring rubric based on key requirements of the RFP. For first-year funding, reviewers include raters both internal and external to IBHE. In the second and third years, funding is based on applications, written in response to a renewal RFP, that include evaluation evidence. Decisions to continue funding are made by IBHE personnel. The overall ITQ program structure is shown in Figure 1.

The Challenge and the Potential Value of an ECB Approach

With discretionary funding so scarce, it becomes extremely important to maximize its impact. How best to maintain flexibility and creativity while increasing expectations for high-impact projects? One answer may lie in increasing the capacity of the state agency and its grantees to embed high-quality evaluation throughout the funded projects and the program as a whole. Evaluation capacity-building, or ECB, is described as follows: “ECB is the intentional work to continuously create and sustain overall organizational processes that make quality evaluation and its uses routine” (Baizerman, Compton, & Stockdill, 2002). In an era of increasing accountability for results, “organizations must develop and nurture a culture of continuous quality improvement that embodies a commitment to ensuring the quality and effectiveness of services” (Ristau, 2001, p. 555).
For the ITQ program, if ongoing evaluation becomes a way of doing business—from the time the Request for Proposals is written until the time the final project reports are filed—it is possible to increase the focus on improving processes and outcomes, potentially achieving stronger impact, not only for individual projects, but also for the program as a whole. This paper draws parallels between evaluation capacity building within funded projects and across the ITQ program as a whole, with the intent of positioning the grantor and the grantees to deliver more “bang for the buck.”

**Theories of Change and Logic Modeling**

Every educational improvement project operates under a “theory of change,” or program theory, representing the beliefs of the project team about how change occurs. Sometimes the theory of change is stated explicitly; sometimes it is implied. Either way, project developers create a series of activities or interventions that they believe will bring about desired outcomes for participants. These activities should lead logically to specific products and effects (Weiss, 1997). If participants do not agree on their theory of change, project success becomes much more difficult (Hatch, 1998).

In ITQ projects, project teams consist of a project director, university and school district personnel, and a project evaluator. A project team creates a logic model, a form of project map, to illustrate the project’s theory of change and make explicit the relationships among the project’s resources and activities and its products and outcomes. *Inputs* are the “ingredients” or resources brought to bear. *Activities* are the actions...
taken or interventions conducted. Outputs are the services and/or products of the activities. Outcomes are the effects, or impacts, of the activities—the desired changes intended by the project (Frechtling, 2007). The logic model visually depicts the project’s theory of change, and each “link” in the model is part of a testable hypothesis and intended results. Project teams create logic models in reverse order. They first conduct a needs assessment to identify weaknesses and determine the desired project outcomes. Next they design the intervention activities, services, and products to be delivered. Finally, they identify the necessary resources and embed evaluation feedback loops. Arrows illustrate logical directional connections among the logic model components. Figure 2 displays the basic logic model structure.

![Figure 2. Basic logic model structure.](image)

Using this basic structure, individual ITQ professional development grants create customized logic models that illustrate their project components and the connections among them. Figure 3 is a logic model that generally applies to most projects funded under the ITQ program, with some project-specific modifications.

As program evaluators, we recommended to state ITQ program administrators that they lead by example and create a logic model and program evaluation plan for the program as a whole. They could consider their funded projects to be their “investment portfolio,” and ask the following questions (adapted from Frechtling, 2011):

1) Does the investment portfolio align with the overall program goals? How has the portfolio changed over time to improve this alignment?

2) Has the state program created and supported constructive and sustainable partnerships?

3) Is the program fostering projects that embed evaluation and gather relevant evidence?

4) Have the projects achieved their goals?

5) Has the program achieved its goals?
State program administrators and program evaluators based the ITQ program logic model on two theories of change: (a) shared knowledge improves practice, and (b) evaluation capacity-building throughout the program and the individual projects will lead to better project and program effects. The first theory involves sharing knowledge about project delivery and evaluation and focuses mainly on continuous improvement. The second theory predicts that embedding evaluation will improve project and program effects and focuses on outcomes. Overall ITQ program effects are determined by aggregating individual project effects. The logic model depicts the relationship among inputs, activities, outputs, and outcomes for the overall ITQ program and includes feedback loops using evaluation data for continuous improvement. The ITQ logic model has two “threads,” based on its two explicit theories of change: (a) the theory that shared knowledge improves practice (upper white thread), and (b) the theory that evaluation capacity-building will lead to better project and program effects (lower gray thread). These theories represent the state agency’s understanding of how improvements will take place.
Figure 4. ITQ program logic model.
Inputs

The ITQ program evaluation and improvement effort is fueled by four “ingredients” or inputs: (a) a multi-year study of funded projects by external program evaluators; (b) a meta-analysis compiled by program evaluators of individual project evaluations; (c) the knowledge and experience of both program evaluators and IBHE program officers; and (d) national standards for program evaluation [Joint Committee on Standards and Educational Evaluation (JCSEE), 1994]. These inputs provide the resources that enable evaluation and improvement activities to occur.

Theory of Change 1: Shared Knowledge Improving Practice (Upper White Thread)

Viable theories of change are supported by the research literature. Research supports the theory that shared knowledge improves practice. This is a well-established premise in business and management, where studies show that boundary-spanning social networks facilitate knowledge-sharing across organizations, creating opportunities for learning and flexibility that would not exist inside a closed organization. This learning and flexibility then provides operational and strategic advantages to organizations (e.g., Nonaka, & Takeuchi, 1995; Liebeskind, Oliver, Zucker, & Brewer, 1996; Brown & Duguid, 2001). Following the theory of change for this thread, the expected outcome is that individual projects and the program as a whole will experience continuous improvement in their operations and evaluation processes, leading to better outcomes for teacher and student learning.

Two key activities enabled grantees, external program evaluators, and IBHE staff to share knowledge. The first was an annual symposium bringing together all parties to discuss the multi-year study, the meta-analysis of project evaluations, the national evaluation standards, and implications for individual projects. The symposia were structured to enable dialog among project teams, who share progress, challenges and promising practices. As an output for this activity, symposium Proceedings were formulated, published and shared among participants. The Proceedings provided a formal record of the materials, presentations, and dialog.

The second knowledge-sharing activity was continuing dialog and feedback and support for individual projects provided by both program administrators and program evaluators. The program evaluation team had become knowledgeable about the projects through multiple site visits, document and artifact analysis, and dialog with project teams, including project evaluators. The program evaluators offered suggestions for project teams to consider regarding implementation and project evaluation. In turn, project teams provided feedback regarding progress, challenges and accomplishments. Outputs for knowledge-sharing activities were net-
working and collaborative problem-solving among project teams, external program evaluators and IBHE staff.

**Theory of Change 2: Evaluation Capacity-Building Improving Project and Program Effects (Lower Gray Thread)**

Since 2003, the evaluation team from the Center for the Study of Education Policy (CSEP) at Illinois State University has partnered with the Illinois Board of Higher Education (IBHE) to improve its grant evaluation system. In an analysis of the 2002 Higher Education Cooperation Act (HECA) grant evaluations, the team found that grant evaluations were neglected by both the state agency and individual grant directors with only 41% of grants submitting final evaluations. In addition, of the 41% that did submit a final evaluation, those evaluations primarily reported process evaluation data related to inputs and outputs. For example, if the proposal promised to offer three summer bus trips to facilitate minority student transfer to four-year universities, then the evaluation simply indicated travel dates and a head count of travelers, the project’s outputs. Very few reported data related to outcomes or results, such as how many minority students from city and community colleges actually matriculated to four-year institutions, let alone whether any matriculations could be attributed to HECA project activities. Even after as many as 20 years of ongoing funding, no evidence of program results was available. With this neglect of evaluation, the IBHE was unable to make either informed funding or other policy decisions, or to report improvements to the education system as a result of the HECA grant funding. Upon reflecting on this problem and the purpose of evaluation in grant programs, the CSEP evaluation team reasoned that grant funding streams could be more than a loose collection of projects. With stronger evaluations serving as action research endeavors within the projects, the evaluation data could be used by IBHE to sharpen policy and add to the research base on effective education programs and practices.

In 2004, the evaluation team had their first opportunity to test their ideas about grant evaluations when the team became the external evaluators of the Improving Teacher Quality (ITQ) Grants awarded competitively under *No Child Left Behind*, Title II. The team began an intentional effort to build evaluation capacity with IBHE staff who administer these grants and the ITQ grantees.

In general, the CSEP evaluation team has been intentionally engaged in evaluation capacity building (ECB) efforts with IBHE and ITQ grantees to embed evaluation, particularly program theory evaluation and logic modeling, as an essential, valuable, and routine project component and as a tool for continuously improving the project. For IBHE, ECB became an approach to managing the ITQ program and assessing its effectiveness in improving teacher quality in projects across the state. In this
section, we have written about the CSEP team’s ECB efforts with IBHE staff, and other chapters in this special issue of *Planning and Changing* have addressed ECB efforts with project grantees (Baker, 2011; Gardner, 2011a, 2011b; Leslie, 2011; Prusaczyk & Baker, 2011; Voss, Khazaeli, Eder, & Gardner, 2011) and present the CSEP evaluation designs and methods (Gardner, 2011b).

Preskill and Boyle (2008) presented a model for ECB that outlined three objectives of ECB in three domains: knowledge, skills, and affect (i.e., beliefs). With IBHE, the CSEP team’s ECB efforts were primarily targeted in the knowledge and affective domains. IBHE staff members are not involved in implementing evaluation studies in each of the projects, therefore the team did not focus on teaching IBHE the skills in implementing evaluation. The team believed that it was more valuable for IBHE staff members to understand the concepts of credible, valid, and relevant evaluation to inform their management of the ITQ grant program. Thus, in the knowledge domain, the CSEP evaluation team has consulted with IBHE staff to educate them about evaluation terms and concepts, how evaluation contributes to decision-making, and the relationship among a project’s goals, objectives, activities, and expected outcomes using each project’s logic model to guide the evaluation process. In the affective domain, IBHE staff learned about the uses of evaluation in producing data to inform and continuously improve programs, thus contributing to the project’s effectiveness. Therefore, evaluation is worth the time and expense as part of the program design and funding. Starting in 2004, the CSEP evaluators and IBHE grant administrators have held three to six regular meetings annually, mutually organized grant wide symposia, set goals for evaluation and policy development and refinement, and supported creation of materials to communicate project findings to IBHE’s trustee board (Gardner, 2011b).

**Feedback Loops**

Both threads of the ITQ program logic model include feedback loops. In the shared knowledge thread, once continuous improvement in delivery and evaluation outcome was assessed, the information (whether delivery and evaluation improved and how much) was used by program evaluators and IBHE staff for continuing discussion, problem-solving, and policy improvements. In the evaluation capacity building thread, once overall program effects were measured, results cycled back to IBHE staff to again refine the Request for Proposals and/or seek other policy methods to improve results.

**Results**

Over the seven years of this ECB effort, the ITQ program has gone from funding 26 projects with marginal to ineffective evaluation plans to
funding nine projects with solid, research-based logic models and aligned evaluation plans and strong evidence of portfolio alignment. Five of these nine are now selected for performance funding and a CSEP meta-evaluation that will be completed in 2011–2012. In 2004, the RFP sought innovation in design and sensitivity to local or regional conditions. But as ITQ matured, IBHE wanted to see results that indicated: (a) full and faithful implementation of professional development plans; (b) increasingly sophisticated project level evaluation used to inform each renewal application or proposal; (c) the viability and sustainability of school-university partnerships; and (d) development of evaluation capacity to link teacher and student learning at the project level. These requirements for pulling together overall impacts in a statewide meta-evaluation are now balanced with the ongoing need for innovation, improvement, and local variation. This special issue of Planning and Changing includes three case studies (Leslie, 2011; Prusaczyk & Baker, 2011; Voss et al., 2011) and articles on project-wide meta-evaluation that share results insofar as possible (Baker, 2011; Gardner, 2011a). Each project is different and not all are amenable to readily comparable measures such as state test scores, so we will not report them here.

There are other complications. In STEM, math testing does not align well with the constructivist teaching practices used in several of the projects (Leslie, 2011; Prusaczyk & Baker, 2011). Science is tested infrequently by the state at grades 4, 8, and 11, so projects must develop their own ongoing measures and help teachers to do their own assessments to guide them (Voss et al., 2011). Projects do use high stakes test results and other measures as a part of their own ECB. Individual case studies included in the special issue and elsewhere offer project results. Each project has its own methods of evaluating implementation fidelity, for example. This has resulted in the development of project specific protocols (Leslie, 2011) and use of research-based implementation measures, like the Surveys of Enacted Curriculum (Wisconsin Center for Education Research, 2011). Here, we discuss our results in terms of what has been accomplished through ECB efforts at the level of statewide policy.

One result of the evaluation team’s ECB efforts occurred when IBHE embedded evaluation within the funded projects. This involved IBHE revising its ITQ Request for Proposals (RFP) to require four new components: (a) an explicit statement of a project’s program theory (theory of change), its research basis, and its visual representation through a logic model; (b) a needs assessment for the first of any three year funding cycle; (c) evaluation plans that measure each logic model as a set of “links;” and (d) evaluation plans that meet national JCSEE standards.

Revising the RFP created new products for the next funding cycle: project proposals with explicit theories of change and evaluation plans aligned with those theories. This change allowed individual project evaluations to measure the intended products and outcomes of each project with more validity and reliability. The feedback provided by the evaluations was
more likely to result in project adjustments that continuously improved the likelihood of attaining the desired project outcomes, allowing for strategic revisions of theories of change. Individual project evaluators did not have to produce a myriad of unconnected measures and reports; instead, they used a streamlined array of measures that specifically gauged progress toward and attainment of the desired outcomes, including explicit measures for teacher learning, classroom implementation, and student learning. For the ITQ program, this adjustment produced an important outcome: more valid and reliable documentation of program effects, along with better understanding of how specific activities produced those effects. Over time, teaching and learning outcomes have improved for individual projects.

It has now become feasible to institute a new policy iteration, a performance funding approach, in which renewal applications that show valid documentation of significant effects are eligible to receive additional funds. In 2011, five of nine projects showed such effects and received performance funding, including all three cases in this special issue. Central to this thread is the understanding that a Request for Proposals is essentially a policy document. The RFP sets forth the state’s interpretation of federal guidelines governing the ESEA Title II grants, along with additional state-specific requirements for implementation and outcomes. The policy guidance set forth in the RFP is intended to make positive project impact more likely. The RFP thus becomes part of the loop of “evidence-based policymaking” (Sanderson, 2002), in which evaluation informs the continuous improvement of policy.

Also central to this thread is the recognition that, in addition to measuring the impact (outcomes) of the individually funded projects, it is also possible to measure the impact of the ITQ program overall. Sometimes program reporting takes the form of numbers: grants awarded, dollars spent, teachers trained, institutions participating. These are good measures of activity (outputs), but are not measures of impact (outcomes). When we look for valid and reliable evidence of program effects (the lower thread outcome), we are measuring the actual impact of state policy on schools, teachers and students. “Policy impact is not the same as policy output. It is important not to measure benefits in terms of government activity…. We cannot be content with measuring how many times a bird flaps his wings; we must assess how far the bird has flown” (Dye, 1972, p. 292). Through strengthening individual project evaluations, as well as evaluating the overall program, the validity and reliability of effect measures can be improved and actual outcomes (“how far the bird has flown”) assessed.

**Evaluation in Action**

In order for IBHE to determine how its overall logic model was operating, the program evaluation team asked questions specifically geared to assess activities, outputs, outcomes, and the links among them. Evalu-
ation questions tested the links in the model. For each theory of change, some of the evaluation questions the team used are listed in Table 1.

Table 1

Evaluation Questions to Test Logic Model Predictions

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does the “investment portfolio” align with the overall program outcomes? How has the portfolio changed over time to improve this alignment?</td>
</tr>
<tr>
<td>• How has the portfolio of funded projects changed over time to reflect the program outcome of continuous improvement in program delivery and evaluation methods?</td>
</tr>
<tr>
<td>• To what extent does the portfolio of funded projects reflect valid documentation of project effects?</td>
</tr>
<tr>
<td>• To what extent does the portfolio focus on increased teacher and student learning?</td>
</tr>
<tr>
<td>Is the state program creating and supporting viable and sustainable partnerships?</td>
</tr>
<tr>
<td>• To what extent are the partnership structures and processes adding value to the funded projects?</td>
</tr>
<tr>
<td>• To what extent are the project partnerships establishing long-term relationships and planning for sustainability beyond the grant period?</td>
</tr>
<tr>
<td>Is the program fostering projects that embed evaluation and gather relevant evidence?</td>
</tr>
<tr>
<td>• To what extent have project evaluations evolved to better measure and report teacher, classroom, and student outcomes?</td>
</tr>
<tr>
<td>• To what extent have non-aligned evaluation activities been reduced or eliminated?</td>
</tr>
<tr>
<td>• How is evaluation data being used to improve individual project delivery and outcomes?</td>
</tr>
<tr>
<td>Have the individual projects and the overall ITQ program achieved intended goals?</td>
</tr>
<tr>
<td>• To what extent do individual projects provide credible evidence of teacher, classroom, and student effects?</td>
</tr>
<tr>
<td>• To what extent does a meta-evaluation of project effects indicate that the program is achieving its goals?</td>
</tr>
</tbody>
</table>

These are occasionally refined to accommodate state and federal reporting requirements, IBHE staff and board priorities, and the knowledge gained from each successive meta-evaluation. At the annual symposia, these questions and the answers gathered by the evaluators helped to guide the structure of large group presentations and small group breakout sessions.

At the time of this writing, program evaluators have accumulated convincing data that the program is better aligning its portfolio with intended outcomes, and that it is building evaluation capacity within its funded projects. With the tighter alignment and focus generated through the most recent RFP, funded projects are entering their second year of activity. We anticipate partnership and outcome data will be available for analysis by the end of summer, 2012.
A State Agency Leading by Example

In government programs, the potential exists for a disconnect between the way the program operates and performs at the state or federal level, and the expectations for operation and performance at the local level. Government agencies, at their best, model what they know to be good practice, holding themselves to the same rigorous standards they expect from grantees and/or contractors. When an agency awards a grant, it is entering into a form of contract with the grantee. As contracts lay out expectations for both parties, so do grant award agreements. These documents are often perfunctory, listing the obligations of both parties to perform the required activities and properly handle the money. When grant project agreements are relegated to a mere compliance obligation, an opportunity is missed to elicit active learning and continuous improvement in project outcomes. If instead, the funding agreements are a mutual commitment to focus on outcomes, implement rigorous evaluations, strive for continuous improvement, and share knowledge, the potential exists for realizing tangible, measurable benefits for project participants. With an evaluation capacity-building approach, a relatively small amount of discretionary funding can indeed go a long way toward creating strong positive effects for teachers, students, faculty, and agency staff.

References


Lynne Haeffele is a Research Associate in the Center for the Study of Education Policy at Illinois State University, Normal, Illinois.

Lisa Hood is a Research Associate in the Center for the Study of Education Policy at Illinois State University, Normal, Illinois.

Bridget Feldman is a doctoral candidate in the Department of Educational Administration and Foundations at Illinois State University, Normal, Illinois.