WCER Working Paper No. 2010-1
January 2010

Designing and Testing Evaluation Tools to Inform Foundation Education Reforms and Build Grantee Capacity

Steven M. Kimball
Wisconsin Center for Education Research
University of Wisconsin–Madison
skimball@wisc.edu

Rachel Lander
Wisconsin Center for Education Research
University of Wisconsin–Madison
rlander@wisc.edu

Christopher A. Thorn
Wisconsin Center for Education Research
University of Wisconsin–Madison
cathorn@wisc.edu

***With* Katie Herrem**

WCER
Wisconsin Center for Education Research
School of Education • University of Wisconsin–Madison • http://www.wcer.wisc.edu/

WCER working papers are available on the Internet at http://www.wcer.wisc.edu/publications/workingPapers/index.php. Recommended citation:

Kimball, S. M., Lander, R., & Thorn, C. A. (with Herrem, K.). (2010). *Designing and testing evaluation tools to inform foundation education reforms and build grantee capacity* (WCER Working Paper No. 2010-1). Retrieved from University of Wisconsin–Madison, Wisconsin Center for Education Research Web site: http://www.wcer.wisc.edu/publications/workingPapers/papers.php

The research reported in this paper was supported by the Searle Funds at The Chicago Community Trust and by the Wisconsin Center for Education Research, School of Education, University of Wisconsin–Madison. Any opinions, findings, or conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies, WCER, or cooperating institutions.

Designing and Testing Evaluation Tools to Inform Foundation Education Reforms and Build Grantee Capacity

Steven M. Kimball, Rachel Lander, and Christopher A. Thorn
***With* Katie Herrem**

Foundation Evaluation Practices: The Need for Better Tools

Foundations use different approaches to evaluate their programs, depending on their evaluation needs. Some foundations merely require grantees to provide a financial report, while others require formal external evaluations (Kramer & Bikel, 2004). Few foundations regularly engage in on-going evaluation activities, and despite the importance of evaluation to philanthropic organizations, "…it seems that the vast majority of the more than 60,000 foundations in the United States do not conduct formal evaluations at all" (Kramer & Bikel, 2004 p. 54). A similar finding also emerged from a recent study of philanthropic organizations in the United States; many foundations see evaluation as costly, confusing, and impractical (Kramer, Graves, Hirschhorn, & Fiske, 2007).

Although practices and definitions may vary, foundations do recognize the importance of evaluation. Strategic planning, decision making, and issue identification are among the primary foundation processes that can be influenced by evaluation. Ongoing evaluations can allow foundations to trace efforts over time, learn from successes and failures, and draw evaluators and foundation stakeholders into a productive dialog about issues and priorities (e.g., Condliffe Lagemann & de Forest, 2007; Kramer & Bickel, 2004; Kramer et al., 2007). Resulting knowledge about both successes and failures can enhance the foundation's own efforts and the efforts of the funding community as a whole. In addition to improving strategic planning, evaluation can enhance accountability, allowing foundation staff and governing boards to respond to institutional funders or individual donors who need to know that money was spent as intended, and that the resulting outcomes justify future efforts. Accountability also stretches to the grant beneficiaries and to the community in which foundations focus their efforts (Kramer & Bickel, 2004).

Recent Trends in Evaluation

Recent trends suggest foundations are using more forward-looking, proactive approaches toward evaluation and assessment. These approaches recognize the potential of formative evaluation to inform both foundation efforts and those of grantees (Kramer et al., 2007). Other new approaches to evaluation are driven by initiatives to make a social impact by building grantee capacity, developing public-private partnerships, investing in social entrepreneurs, and funding advocacy efforts to shape public opinion.

This paper was previously presented at the annual conference of the American Evaluation Association in Denver, Colorado, November 6, 2008. The authors express their sincere appreciation to the education program officers of the Trust for embarking on this pilot study and for the time invested in this project by the Big Shoulders Fund and the teachers and leaders of its member schools. Correspondence may be sent to: Steven Kimball, 1225 West Johnson Street, Suite 653, Wisconsin Center for Education Research, U.W.-Madison, Madison WI 53706, or by email at: skimball@wisc.edu

It appears that organizations are moving away from traditional evaluation approaches that look to the past in order to retrace what has occurred. Instead, they increasingly use performance-based and utilization-focused evaluation tools that can improve grants or programs as they unfold (Patton, 1997). These strategies may reduce the applicability and utility of traditional evaluation designs intended to demonstrate causation (Kramer et al., 2007; Kronley, 2000). Using these new approaches that are tailored to the needs of individual projects makes it hard to compare results across grantees and to assess the realization of foundation strategies. Acquiring relevant information requires a more hands-on, informal approach to grantee evaluation, with more emphasis on conversations and site visits. The increasing use of interim goals among nonprofits and foundations is another indication of the trend towards collecting timely information to inform decision-making and improve program performance (Kramer et al., 2007).

In response to evaluation challenges, several foundations, including the W. K. Kellogg Foundation, Annie E. Casey Foundation, and United Way of America have embarked on long-term and comprehensive efforts to establish evaluation resources for grantees and nonprofit organizations (see, for example, Hendricks, Plantz, & Pritchard, 2008; Behrens & Kelly, 2008; and W. K. Kellogg Foundation, 1998). These resources can help other foundations and organizations who are seeking guidance on where to start and how to focus evaluation efforts. Based on their study of nonprofit organizations in one state, Carman and Fredricks (2008) also assert that foundations are uniquely situated to help support and train nonprofits in evaluation. Despite emerging resources and increased attention to the field of evaluation in nonprofit organizations and foundations, it has been recently argued that, "The field of philanthropy urgently needs better tools to help funders learn from one another and to identify the evaluation approaches best suited to different circumstances" (Kramer et al, 2007, p.6). In this paper, we report on evaluation resources developed for a large community foundation in the Midwest.

Initial Investigation and Development Work

Evaluating The Chicago Community Trust Education Initiative

Beginning in 2002, The Chicago Community Trust embarked on an ambitious grant-making strategy to improve education outcomes primarily in the city of Chicago and Cook County, Illinois. Known as The Education Initiative, this effort focused on three priority grant areas: literacy, professional development, and alternative models of schools. The goals of the initiative were to: (a) develop high performing schools and professional networks of educators to support rigorous instruction that reduces the achievement gap; and (b) increase the number of high quality alternative educational opportunities for students. The 5-year investment strategy ended in September 2006, with over $55 million awarded for 317 grants across the 3 priority areas over the 5-year period. The Trust is currently embarking on a new education program and has pledged an additional $50 million to support educational improvement in the Chicago community over the next 5 years.

In June, 2005 (the 3rd year of the initiative), the Trust awarded a contract to researchers at the University of Wisconsin–Madison, Wisconsin Center for Education Research (WCER) to assess the impact of the Education Initiative. The evaluation team was instructed to provide a comprehensive, 3-year evaluation. Specifically, the Trust wanted to learn about the reach and impact of the Education Initiative on the education system in the city of Chicago, and assess possible current and future changes to their grant-making strategy. The evaluation was designed to provide both ongoing, formative feedback to the program officers and executive committee, and summative information on the impact of the initiative. Formative interim deliverables, as well as year-end reports, were provided to the program officers during each year.

During the 3rd year of the evaluation, the program officers and evaluators were especially interested in pilot testing evaluation guidelines that were developed as part of the overall evaluation of the Education Initiative. These guidelines and related evaluation reporting tools were intended to help grantees tailor projects to the foundation's priorities, and help the foundation assess the impact of their diverse grant activities.

We report here on the design and pilot testing of the evaluation guidelines and reporting tools we developed to support the education program of the Trust. The next section describes the impetus for developing these tools and the evaluation findings that informed their design. In the fourth section, we briefly describe the pilot study design and the participants. We also report findings from the pilot study. In the final section, we describe implications for use of the guidelines and tools by the Trust and other foundations.

Existing Evaluation Practices at the Trust

Issues with Primary Data Sources

We initially planned to use grant documents, including proposals, evaluations, and final reports as primary data sources. However, the program officers noted at the start of the evaluation that these reports varied in quality and were probably not suitable for a meta-analysis approach. We would also have to pursue original sources of data. Nonetheless, given our time and resource constraints, the grant documentation was a primary evidence base for the 1st year evaluation report. The analysis started with the grant evaluation requirements.

In the early years of the Education Initiative, grantees were only required to provide a final report that included a narrative of the project and a financial statement. There was no standard format for reporting these elements. Although the Trust staff engaged in planning retreats to focus on evaluation, during the early years of the initiative attempts to standardize evaluation reporting were seen as too constraining for projects with different foci. During 2004, a new online grant proposal format was introduced across programs at the foundation, which included prompts for expected impact, outcomes, and measures.

Revised Reporting Requirements

The 2004 requirements, still in use, called on prospective grantees to include an evaluation plan as part of their grant proposals. The evaluation plan requirements included a short overview of the anticipated evaluation focus, followed by the expected impact of the program, with up to three outcomes and one to three measures to assess each outcome. No guidance was offered on these components. The prompts were general and left discretion to grantees to determine appropriate evaluation activities. However, program officers did react to proposed evaluation plans as part of their award approval process. Many grantees called on the program officers to ask for guidance on what they expected to see in the evaluation reports at the end of the grant.

It is perhaps not surprising that grantees' reports and evaluations varied widely in depth, quality, and accuracy. As one program officer noted, "The mixed bag of reports during these early years on all the little projects reflected, I think, a period of multiple and constant changes in the Program Department [and] the Trust as a whole." Problems with reports included a lack of information on evaluation processes, instrumentation, outcomes, and lessons learned. In some cases, the focus of evaluation was not appropriate for the stage in which the grant was being initiated. Although there were examples of strong and influential external evaluations of larger projects, the evaluation reports from many smaller projects were inconsistent, incomplete, or missing. It was difficult to identify through grant documentation maintained by the Trust evidence that could support grantee claims of impact, effectiveness, or lessons learned. As a result, assessing the impact of the overall grant-making strategy was hampered.

Evaluation Practices of Peer Foundations

Foundation Interviews

Although the literature referenced in the introduction indicated some general trends in foundation evaluation, in order to gain more specific insight into current evaluation practices, we conducted telephone interviews in the spring of 2007 with a peer group that included seven international, national, and community foundations across different regions of the United States. Organizations were selected in consultation with the Trust education program officers to gain an understanding of the range of evaluation and grant reporting practices that exist at different organizational sizes, giving levels, geographic regions, and funding emphases. From each of these foundations, we spoke with individuals in senior positions who had direct experience with or supervision of their organization's grant evaluation practices. Table 1 includes an overview of the foundations sampled.

Table 1
Foundations Included in Study With Selected Demographic Information

Foundation type	Location	Assets	Annual giving
International	East	$2.5 billion	About $90 million (FY 2006–2007)
Community	Midwest	$1.9 billion	About $84 million (2006)
International	West	$33.4 billion	About $1.56 billion
Community	Midwest	$1 billion	About $172 million (2006)
Regional	Midwest	$900 million	$50 million (2006)
National	East	$9.4 billion	$403 million (2006)
Community/ Regional	West	$1.7 billion	NA

We interviewed these foundations to learn about their current evaluation requirements and practices. Specifically, we wanted to understand (a) what types of evaluation plans their applicants must submit, (b) if they provide any evaluation guidelines or indicators for their grantees, and (c) what level and type of staff and data tracking infrastructure they use to support their evaluation activities. We also asked questions to understand if and how they synthesize their evaluation findings across grants to provide a picture of initiative- or program-wide accomplishments. The study provided additional detail about evaluation considerations among foundations as well as some interesting developments in evaluation activities.

Evaluation and reporting requirements. Nearly all foundations in our sample have evaluation and reporting requirements for their grantees, but their requirements vary by scope and formality. In most cases, potential grantees are asked to describe their evaluation activities in the grant proposal. The extent to which program officers or other foundation staff influence evaluation plans or monitor evaluation and grant activities often depends on the size or the potential impact of the grant.

For larger grants or those perceived to have strong potential to inform future efforts, the foundations may require an external evaluator. In one example, a foundation used several different evaluators applying different evaluation approaches to better understand a complex program initiative. Another foundation requires external evaluations for most of their major, national programs. The regular use of external evaluators is not the norm for most grants.

For smaller projects, many of the foundations simply require the grant applicant to specify some program outcomes or a general evaluation plan, and accept periodic reports in lieu of formal evaluations. In several cases, such monitoring is done via informal check-ins with grantees. Only one of the respondents indicated that evaluation plans were not required as part of the foundation's proposal process; instead, they use a standard, basic reporting format to obtain semi-annual reports from grantees, then follow up with informal but frequent contact between the foundation and the grantees.

One notable approach to ensuring the quality of grantee evaluation is one foundation's practice of verifying grantee reported data by collecting data from other sources. According to this organization's representative, this approach points to the importance of creating appropriate evaluation measures. An example was provided of a grantee reporting gains in test scores due to a funded intervention, but the foundation's check of student grades indicated little progress had been made.

Several comments mentioned the importance of achieving a balance between the need to gather useful evaluation data and the burden evaluation and reporting requirements may place on both grantees and foundation staff. One respondent emphasized the importance of foundations working directly with grantees to create a report format and evaluation criteria that is beneficial not only to the foundation, but also to the grantee. This person also commented that the approach is recommended for foundations working with grantees that are in the monitoring and reporting phase, in contrast to grantees conducting larger projects that have external evaluators.

Use of guidelines. The use of specific evaluation guidelines is uncommon among the foundations. Most reported that evaluation indicators and approaches are tailored to each grant and its unique goals. In some cases, grantees work directly with the foundation to create appropriate evaluation measures, and the indicators are built into the evaluation plan in the proposal stage. For external evaluations, the external evaluator may have a large say in how evaluations are designed, with varying input from foundation staff. In other cases, grantees are largely left on their own to develop, implement and report on their evaluation activities.

Rather than employing evaluation guidelines for all grants, it was more common for the foundations to require grantees to provide written reports on their activities either annually or semiannually. In some cases, common reporting forms are used. However, several of the respondents admitted that the reports are more of a formality, and while they are routed to the appropriate program officers, the reports do not necessarily influence the program or foundation strategy. One foundation did use a structured format at one time for all evaluations. Their approach was based on a logic model for reporting on grant activities and outcomes. The foundation dropped that approach and reverted to collecting more grant-specific narratives from grantees on what was accomplished. It was unclear from this respondent, however, how potentially idiosyncratic evaluation approaches and reports were verified or used to inform future foundation activities. Another foundation that has been highlighted in national publications for their strong evaluation strategy prefers using logic models rather than less uniform evaluation approaches. Their strategy, which is described below, helps them synthesize results across major funded programs.

One of the ideas most commonly articulated by interviewees was the importance of obtaining clear information from grantees up front. As one respondent reported, “You have to get to the specifics of what is the result at the end of the day. If you don’t know what you want on the front end, it ends up as decent charity at best.” Many of those interviewed similarly indicated that they wanted clear ideas from the start about the goals for funded programs, what programs intend to produce, and how success will be measured.

Infrastructure to support evaluation. Respondents reported a wide a variety of approaches to sustain their grant evaluation activities. One of the larger foundations employs staff members who are experts in evaluation and assessment; each plays a major role in the design of grant evaluations. Other foundations grapple with the issue of providing adequate staff to oversee evaluation activities and rely on program officers to ensure that evaluation activities are completed. In some instances, foundations don’t have enough staff to maintain frequent contact with grantees. Data-tracking systems also vary greatly among the foundations. One foundation is preparing to roll out a new data-tracking program that will maintain detailed records in common measures across several education-related grants. Another foundation does not have a database for tracking grant results or impacts and merely maintains paper records.

Synthesizing results. Most of the respondents articulated a desire to use grant evaluations and cross-project evaluation synthesis to be more strategic in grantmaking efforts. Their organizations have struggled with assessing the impact of multiple grants, both within programs and across program areas. The challenge is compounded since evaluations were often tailored specifically to individual grants, if they were conducted at all. Yet, many were seeking ways to maximize their strategic influence or to hold their organizations internally accountable for results. They saw the benefit of identifying common performance indicators that allowed for systematic analysis of grants, even when their current practices worked against these goals due to their individualistic nature.

Despite the difficulty in institutionalizing a process for synthesizing results, several foundations articulated examples of attempts to synthesize evaluation findings from multiple grants. These practices are highlighted below:

- One foundation reported that program staff provides quarterly reports to the foundation board on what they are trying to achieve in their program areas. They attempt to aggregate grantmaking impacts and the foundation's influence as a whole. This effort includes collecting and analyzing secondary data to confirm or refute grantee reports and external economic and community indicators of foundation impact, such as growth in a particular industry of interest to the foundation. An external evaluator is also called upon annually to provide them with a landscape of the foundation's work, which includes examining the institutions the foundation has invested in and how they are performing.

- Similarly, another organization attempts to synthesize findings by having program officers complete a closing assessment on each grant. The program officer assesses the grant, draws conclusions about its impact, and communicates lessons learned.

- Another foundation uses a system of logic mapping and path analysis to synthesize grant evaluation findings at the strategy level. This synthesis is part of a five-level foundation evaluation approach that also includes grant-level reporting, external evaluations, and a foundation-wide balanced scorecard. In order to evaluate their grantmaking on the strategy level, this organization hypothesizes how they plan to bring about change within a certain strategy area and creates indicators that will measure that change. As grants are made and evaluation reports provided, the foundation measures results against the logic map, or indicators, to determine what progress is being made and what changes might need to happen next. According to this respondent, the logic maps serve as a good management tool, answering the questions of: "Are things working the way we thought they were going to work? Do we need to up the dose?"

- Another innovative practice, still in its pilot stage in one foundation, is the use of software to collect real-time data in common terms across grantees, programs, and geographic locations. Historically, this organization used inconsistent data collection efforts and varying evaluation techniques, which prevented them from obtaining a comprehensive picture of school and grantee performance. The foundation is introducing a software program that enables them to collect school-level data from grantees and schools using a common taxonomy and data-management system in order to analyze grant outcomes in a consistent manner. At the same time, this foundation is creating a new evaluation framework. The new performance measurement strategies they will employ include a focus on both program implementation and impact across geographic locations and education grantmaking priorities. Using specific "learning questions" to evaluate program progress, this foundation seeks evaluation results that will assist them in refining their strategies, increase grantmaking impact, and provide models of success to showcase and replicate. The ultimate goal is to integrate the new data-management system with complementing performance-measurement strategies and systems to create a holistic approach to evaluation and performance management.

Summary

The foundations we interviewed all place importance on evaluating their grantees' work, and most are actively thinking about or developing systems to improve their grant-making strategies. However, at this time, most of these organizations do not have a specific evaluation framework that is applied to all grants, but rather they employ an ad hoc approach, tailoring evaluation plans to each grant or a cluster of grants. This practice frustrates the desires of many of these organizations to be better able to synthesize evaluation findings across grants and use the evaluation findings to inform strategic planning and decision making.

Developing Comprehensive Evaluation Guidelines and Reporting Tools for the Trust

Framework and Model

After analyzing current literature on foundation evaluation and the evaluation practices of peer foundations, we worked with the education program officers to develop a set of user-friendly evaluation guidelines based on a Lifecycle and CIPP model. The CIPP (Context, Input, Process and Product) model of evaluation, developed by Daniel Stufflebeam and colleagues (Stufflebeam, Madus, & Kellaghan, 2000), emphasizes the systematic acquisition of information that is most relevant for program management and evaluation. Evaluations correspond to decision-making needs, and can adapt to changing needs where necessary. The Lifecycle/CIPP model specifically targets evaluations that encompass the full life cycle of a project, in this case from grant application through final reports. We selected the CIPP framework to address the broad range of questions requested in the original evaluation solicitation and help frame our 1st-year evaluation report. The Trust wanted to learn about grants that varied in scope (from planning to large scale) and content (from charter school improvement to school literacy development). The CIPP model, combined with a life-cycle conception, allowed the evaluation to encompass the life trajectory of grants and yield feedback at all stages of implementation, an issue also discussed in Raymond (2001). Figure 1 illustrates the Project Lifecycle/Evaluation model that forms the basis for the evaluation guidelines.[1]

[1] We noted that other tools are available for use by nonprofits and foundations in evaluating their programs. Some of these methods include Outcome Measurement (e.g., United Way, Outcome Measurement Resource Network, http://national.unitedway.org/outcomes/) and evaluation checklists, such as Key Evaluation Checklist (e.g., Scriven, www.wmich.edu/evalctr/checklists) and Utilization-Focused Evaluation Checklist (i.e., Patton, www.wmich.edu/evalctr/checklists). While these and other frameworks take different approaches to program evaluation, they have several key attributes in common that can assist foundations as they explore new ways to evaluate their programs. Each framework seeks to elicit an understanding of the program's context, stakeholders, activities, products, and outcomes. These general components lead the evaluator towards synthesizing and disseminating program evaluation results in a manner that incorporates evaluation findings into ongoing program adjustment and improvement. Regardless of the framework used, foundations, grant recipients and evaluators are also encouraged to reference the Program Evaluation Standards of the Joint Committee on Standards for Educational Evaluation for evaluation planning and reporting and for meta-evaluations (see http://www.jcsee.org/program-evaluation-standards).

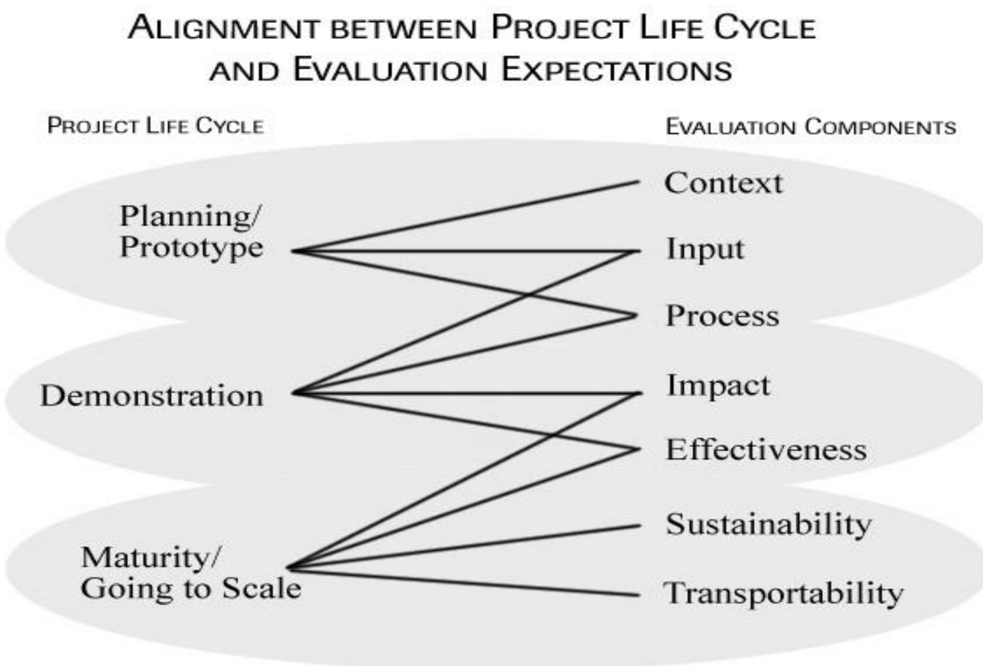


***Figure 1*. Illustration of alignment between project life cycle conception and key evaluation components from CIPP model.**

This model aligns the stage at which a program is beginning (lifecycle) with the most appropriate evaluation questions for that stage. The framework allows for feedback at all stages of a program's implementation, and sheds light on issues surrounding sustainability, transferability, and institutionalization. The guidelines were designed for use by program officers and future grantees to inform and direct grant evaluation activities across the education program priority areas.

The CIPP evaluation framework can be conceptualized as relating to key features of the organizational life cycle. Such a framework can help fill the need for a flexible evaluation strategy that makes sense to grantees, helps inform program planning, and generates comparable data to measure program impact and inform strategy development. The draft evaluation guidelines that we developed for the Trust (see Appendix A) were designed to address a number of issues uncovered in our evaluation reports and the literature about foundations and evaluation. We argued that the use of the guidelines by grantees could yield a variety of benefits for grantees and the Trust, as well as other foundations, including the following:

- Evaluations based on lifecycle meet the multiple needs of projects of varying size and scope.

  The Chicago Community Trust Education Initiative has funded and continues to fund a variety of projects and programs whose impacts extend to the Chicago Public School system and beyond. A primary challenge exists in evaluating such grants, especially assessing impacts across different types of grants. For example, multi-year grant programs receiving large amounts of funding are more likely to want and be able to produce evaluations that assess the program's impact, transportability, and sustainability beyond the funding period. These types of programs may seek to improve student achievement or district practices on a large scale. By contrast, small pilot programs rarely have the capacity to evaluate these effects, and more appropriate measures of program success include an evaluation of the context of the program and the process by which it was implemented in its initial stages.

By employing a common set of basic evaluation questions appropriate to the lifecycle stage of each grant, grantees have a framework upon which to base their evaluations that will provide them stage-appropriate findings to use in program planning, communicating results to beneficiaries and stakeholders, and reporting grant processes and impacts. Furthermore, by using a consistent evaluation method, if it so chooses, the Trust can more easily track grantee evaluations and reports and assess the extent to which each program meets its intended goals and outcomes. Using a common set of guidelines based on project lifecycle lets both program officers and grantees know what types of evaluation questions are appropriate for which grants, reducing the potential for confusion, complication, and inappropriate, inaccurate, or ineffective evaluations.

- Evaluation requirements are clear from the beginning.

  By establishing effective evaluation plans as part of the grant proposal process, grantees and program officers can work together using common language and expectations to formulate program goals, associated measurable outcomes, and impacts that are agreeable to both parties. Grantees will learn what types of questions are appropriate to ask when evaluating their programs, enabling them to learn and report only the most relevant findings to the Trust. Ultimately, this system can help reduce the potential evaluation burden on grantees by articulating to them at the earliest stages what is required and expected of their program and their internal evaluations.

- The guidelines enable grantees to work more effectively with external evaluators on larger projects.

  While external consultants would not necessarily be required to adhere to new evaluation guidelines, a grantee's familiarity with the process and terminology of the evaluation framework will enable them to more effectively communicate with external evaluators throughout the program and evaluation processes.

- A common framework facilitates cross-program synthesis of evaluation findings.

  Among most foundations, there is a pervasive difficulty and limited organizational capacity to synthesize evaluation findings across multiple programs. By using a common framework for all evaluations, foundations would be better equipped to evaluate funded programs within and across priority areas.

- The guidelines are conducive to strategic planning and can inform decision-making.

  This attribute is beneficial to both grantees and foundations as they seek to be more cost-effective and strategic in programming and decision-making. The literature and our interviews with foundations indicate an increasing trend towards foundations' desire to be more strategic. By using a common evaluation framework for all grantees, foundations would be better able to assess the true impacts of their grants. Having an accurate report of different approaches and programs can aid in decision making (e.g. to continue current funding efforts or redirect support to more effective approaches) and strategic planning for future grantmaking initiatives (e.g. looking across programs, which approaches have been most successful? What issues should be priorities during a new initiative?).

Likewise, grantees would have a clearer sense of the purpose of program evaluation and have guidance in how best to use evaluation findings to improve their programs and maximize impacts. Ideally, grantees' views of evaluation would shift from feeling a sense of burden and confusion about evaluation requirements to understanding the value program evaluation has for their projects.

Draft Guidelines

In order to gauge initial grantee reactions to the proposed evaluation guidelines, we vetted the new guidelines with two current Trust grantees and three evaluators who had worked with the Trust.

Grantee feedback. In June 2007, we asked leaders from two Trust-funded projects a series of questions about their funded program, evaluation plan(s), communication with the Trust regarding evaluation and reporting requirements, and progress towards completing their grants. We also asked for their comments on the new guidelines document. These grantees were in the process of conducting similar in-house assessments to aid data-driven instruction in their respective schools. At the time of the interviews, the grantees were at different stages in their evaluation processes. One representative reported that their organization had completed a mid-year report for the grant and was nearing completion of their annual report. The other representative indicated that their annual grant report was due in December 2007.

One grantee reported that their group was unclear about the initial expectations for their evaluations (prior to receiving the guidelines), while the other group felt confident in their understanding of what the Trust requires in the way of evaluation. At the time of the interviews, both grantees had made progress in their evaluations. One had received a completed evaluation from an external evaluator and had already begun to use the recommendations to adjust their project.

In response to questions about the guidelines documents, one grantee said that the document is extensive and somewhat intimidating, though the questions in the guidelines would be useful to them. The other grantee liked the detail provided by the guidelines and thought they would be useful, particularly in the sense that it would afford less experienced grantees a broader conceptualization of what program evaluation can be. Both organizations reported that they could envision using the guidelines to inform their reporting to the Trust.

These discussions resulted in some helpful recommendations for making the guidelines more user friendly. Both grantees expressed an interest in seeing more connection between the evaluation questions and a report shell or format. Grantees seemed eager to ensure that their final evaluation product or document met the reporting requirements of the Trust. The connections between the evaluation guidelines and a final report were not evident to them at the time of our interviews. Prior to these interviews, we drafted a report shell to correspond with the guidelines document in which grantees could fill in their evaluation findings under the appropriate headings. The grantees did not see this draft document, which was subsequently refined based on program officer feedback to provide consistency across projects.

Respondents also suggested that it would be helpful to see the evaluation process streamlined and online, such as has been done with the Trust proposal process. Other recommendations included (a) a sample evaluation report based on the new guidelines, (b) a list of recommended external evaluators, and (c) for the Trust to facilitate the sharing of best practices among area schools with similar projects.

Evaluator and Trust feedback. To further refine the evaluation framework, we participated in a meeting in July 2007 with the Trust program officers and area evaluators with whom the Trust has worked. The other evaluators provided valuable feedback, most of which centered on (a) the usability of the document, (b) consistency in the progression of program development and related evaluation questions, and (c) accuracy in communicating the Trust's purpose for requiring grantees to use the guidelines. The other evaluators commented that definitions are a key issue in refining the guidelines. They were concerned that questions should be included to elicit a theory of action for the evaluated program, as well as an assessment of sustainability at different lifecycle stages. After considering the other evaluators' comments, we made several adjustments to the draft guidelines.

Before and after the July meeting, we worked with Trust program officers to draft and edit the guidelines. This process allowed us to improve the evaluation framework as well as tailor the tool to the needs of the Trust and its grantees. The next step in the project was to test the guidelines with a larger group of grantees via a pilot study.

Pilot Study

We developed evaluation guidelines and reporting tools to establish consistency in evaluation reports and to help the Trust assess the impact of its diverse grant activities. The guidelines and tools were also intended to focus grantee evaluation efforts on the most important aspects of their project so that they could assess, learn from, and disseminate their results.

Once the guidelines and tools were developed, we tested them in a pilot study. The pilot study participants included grantees who, in the past, had submitted reports of limited depth and varying quality. These organizations also had limited capacity and resources to dedicate to evaluation activities. In addition to the guidelines included in Appendix A, the pilot project also tested an evaluation plan template (see Appendix B), a basic logic model format (an example of the modified, completed version included in Appendix C), and the evaluation report template (see Appendix D). The pilot project was designed to test these tools and provide professional development on effective evaluation practices for grantees who participated.

Research Questions

The study addressed two primary research questions:

1. How useful are the grant evaluation guidelines and related tools in helping grantees develop, monitor, and report on grant activities?

2. What are the implications for applying the guidelines and tools across the education program?

Project Participants: Big Shoulders Fund

Background

The project participants included the Big Shoulders Fund (BSF; www.bigshouldersfund.org) and its sixteen professional development projects funded for 2007–2008. BSF was initially formed in response to waning student attendance and lack of financial resources to support the schools and their educators. BSF works with private, parochial schools in the City of Chicago that serve large numbers of low income and minority students. The mission of the organization includes developing teaching and leadership within their schools and helping the schools remain open as viable educational alternatives for Chicago families. The organization has received substantial support by the Trust in the past.

BSF provided a useful case to test the evaluation guidelines and tools due to its role as intermediary between grant funding provided by the Trust (and other organizations) and the schools and projects that apply the funding to their work. BSF attempts to maximize the use of the resources and has responsibility for synthesizing results across projects in their reports of grant use and effectiveness. In this sense, it has some similar functions and needs for grant requirements as the Trust.

Big Shoulders provides support to 93 high-risk, inner-city schools within the Archdiocese of Chicago. During 2007–2008 BSF received $500,000 for its professional development support and directed the funding to 16 member schools that successfully applied for funding. Part of the organization's focus is on helping schools with specific professional development, which is a substantial break from the past. Previously, BSF funding to schools was primarily a "pass through" function, with funding applied at school discretion. Encouraged by the Trust, the organization now takes a more active role in working with schools and is more selective in grant making. In particular, the Trust pushed to end one-time workshops, which tend to have minimal impact. In addition, BSF has been deliberately trying to improve its effectiveness as a grant-making organization to its member schools. Not only has it been trying to create a more uniform end-of-grant reporting process, but also it is moving to improve accountability for funding use, and to get project leaders talking about their progress and sharing lessons learned about how to assess and improve their programs. At the same time, however, the organization is cognizant about the resource limitations facing their schools, and strives to limit the burden on school leaders.

Based on feedback from school leaders, the BSF focuses its professional development grant making on (a) helping schools use assessments as a tool, (b) maximize instruction with technology, (c) incorporate the arts, (d) improve content knowledge, and (e) provide executive education for administrators. Appendix E includes descriptive information on the 16 2007–2008

BSF grantees, including project names, funding amounts, stage in the grant lifecycle, and a summary of each project's purpose.

As reflected in Appendix E, the 2007–2008 funded projects included activities that generally fit within three groupings designated by BSF: content/literacy, data to inform instruction, and technology. One project was directed at marketing and fund raising for school leaders and another was designed to develop a coaching network to provide tailored professional development to four BSF schools over a 3-year period. The projects were funded over a range from as low as $5,000 to as high as $88,000. Projects funded at the higher amounts are those that include multiple schools and tend to have a technology focus (i.e., purchasing and training on instructional, interactive whiteboards). None of the projects had funding budgeted specifically for evaluation activities.

Existing Evaluation Practices

The BSF evaluation plan was intended to encourage projects to list outcomes and how they would be measured, but little guidance was provided on specific aspects of expected evaluations. In this way, BSF was reflecting the requirements of the Trust. With their proposals, BSF applicants were asked the following:

> Describe the measurable short-term and/or long-term outcomes of the project. Please explain how this program will deepen educators' knowledge of the subjects they teach. Also, please specifically describe how the project will impact your school, your teachers, and ultimately, your students.
>
> How will the outcomes be measured? When will data be presented? How will the people who are served by the grant be involved in the project evaluation?

These questions do address important aspects of evaluation, including expected outcomes and how they will be measured. They also seek a timeline on data presentation and probes how stakeholders will be involved in the evaluation. Further, they attempt to get applicants to address questions about how the programs will impact educator knowledge and results for school, teacher and students. These components can then contribute to the final report provided by BSF to the Trust on grant activities and outcomes, which is described within the original BSF evaluation plan included in its application to the Trust. A selection from that initial evaluation plan is presented below in Figure 2.

One intention of the BSF evaluation approach is to encourage it's grantees to develop evaluation plans that focused on outcomes. Although not explicit here, BSF wanted evaluation results to produce some data comparable across grants on expected changes the projects would produce in students (i.e., student achievement and attendance), teacher knowledge, and instructional change. Although these are certainly important overall goals, the focus on these common outcomes and measures may not apply to grants according to

> All accepted proposals must include a well-constructed evaluation plan with measurable outcomes. BSF expects schools to include methods for assessing the efficacy of the programs. Such measurable outcomes include an increase in student performance (indicative of implementation of best practices) and student attendance (indicative of student participation/motivation) as well as an increase in teacher retention rates (indicative of satisfaction with the program and buy-in). Big Shoulders will encourage the schools to include in their proposals both a pre- and post-professional development survey for all participating teachers and administrators. Depending on the grant, schools will be expected to administer surveys to parents and students when appropriate. Furthermore, schools will be required to turn in a mid-year report to Big Shoulders by December, describing the progress and challenges of the program to date. In addition, Big Shoulders' staff will conduct site visits and conferences on an as-needed basis with the schools and professional development providers. Upon completion of the school year, a final evaluation report from the schools must be submitted within one month. Schools can be renewed for grant funding provided that accountability for and effectiveness of the particular project has been determined through this final review process. Once the Big Shoulders Fund receives all school reports, it will submit a final comprehensive report to the Trust in the summer.
>
> **Impact:** BSF expects the impact of professional development to be evidenced by increases in student performance/motivation; in teacher knowledge, instructional strategies and satisfaction; and in the administrators' ability to sustain a community of learning.
>
> **Outcome 1:** An anticipated outcome is that students will be more energized and invested in their education as teachers learn new instructional approaches and diversify their lesson plans with, for example, arts and technology.
>
> > **Measure 1a:** The impact of the program will be measured by the improvement in student performance as demonstrated by student assessments and anecdotal evidence from teachers and administrators.
> >
> > **Measure 1b:** The impact will be measured by the improvement in student attendance rates as an indicator of increased motivation and participation and anecdotal evidence from teachers and administrators.
>
> **Outcome 2:** Another expected outcome is an improvement in teacher content knowledge and instructional strategies as well as teacher satisfaction with their profession.
>
> **Measure 2a:** This improvement in the teachers' content and instructional skills will be measured by evaluations conducted by school leadership and professional development providers and by pre and post professional development surveys.
>
> **Measure 2b:** Schools will use the pre and post implementation surveys and the teachers' retention rates as a measure of improvement in teacher satisfaction…

***Figure 2*. Selection from Evaluation Plan for 2007 BSF grant application**

their lifecycle (e.g., a planning grant). Further, as seen in the evaluation questions to grantees reported above, these more specific expected outcomes were not provided in this level of detail to the prospective or funded projects.[2] Finally, there are no prompts in either the evaluation questions for grantees or in the plan submitted to the Trust for how the projects will learn from their activities through the evaluations and make appropriate adjustments.

In prior years, the BSF evaluation report to the Trust did include summaries of key activities of each grant and perceptions of their impact. These results were presented in a consistent summary format. However, it was often unclear how the results were derived through evaluation, including what evaluation questions were addressed, what specific measures were applied, how the data were analyzed, and what implications the findings had for the future of the projects. Indeed, some findings appeared more anecdotal than data based.

Pilot Study Activities

In December 2008, an initial meeting was held between our evaluation team, the Trust's education program officers, the Big Shoulders Fund, and another grantee of the Trust education program.[3] The meeting was convened (a) to brief the participants on the background leading to the pilot study, (b) explain the evaluation materials, (c) outline expected activities and (d) address any questions about expectations or demands. We shared the evaluation materials and briefly explained them. Materials included the evaluation guidelines document, the evaluation plan template, and evaluation report template. Grantees were informed the pilot project was voluntary and that products and outcomes would not be used to determine future grant awards. Both organizations agreed to participate.

Three 90-minute professional development sessions were held by the BSF director of external affairs. The external affairs director was responsible for coordinating grant reviews, awards, and reports, including the final synthesis of the reports submitted to the Trust. The sessions were split according to the clusters of grants determined by BSF. The clusters included five projects involved with curriculum mapping; five working on technology; and four focusing on literacy or other content-related professional development. The other two involved a professional development provider working with a number of BSF schools, and a marketing project directed toward 36 BSF schools. In total, leaders from 16 projects attended the sessions. The meetings were led by the evaluation team, in conjunction with the BSF external affairs director, and focused on using the evaluation guidelines, report plan template, strategic mapping tool, and final report template. We also provided support on developing and executing evaluations that included written and verbal feedback to grantees. In addition to the three meetings for the 2007–2008 projects, an informational session was held by the evaluation team and BSF in the spring for prospective 2008–2009 applicants and a meeting in the late summer for successful applicants.

---

[2] Applicants were required to provide some detail on project need, program elements, implementation plan and timeline, which could help in developing evaluation plans.

[3] Due to space considerations, we are only reporting findings from one of the primary grant organization, the Big Shoulders Fund and its participating sub-grant projects. Findings were similar for the other organization.

During the first meeting with grantees, held in January, we provided an overview on the pilot study, including (a) a brief explanation of the impetus for the study, (b) the intended benefits of the pilot project to the grantees, BSF, and the Trust, and (c) an overview of the evaluation guidelines document. We emphasized that participation in the activities and application of the evaluation resources was voluntary. Following a question and answer session about the guidelines, we reviewed the draft evaluation plan template and strategy map tool. Participants then were asked to try out the strategy map alone and with a partner. The whole group then discussed the results. Participants were asked to use the evaluation plan template to compare their results with what they had originally proposed as part of their grant application, and complete items prompted in the template that may have been missing from their original plan. The evaluation team offered to give feedback on the revised plans once they were completed. Those participants completing and sharing the plan prior to the March meeting were provided written feedback. An example of written feedback to one grantee is included in Appendix F.

The March meeting was organized to first share overall reactions on the revised evaluation plans. All 16 projects submitted plans, but only six included draft logic models with their plans. During this meeting, the evaluation team first explained the development and use of the feedback forms used on the evaluation plans, with emphasis that we were not evaluating each project per se, but trying to highlight aspects of the plans that looked promising and provide suggestions in other areas that could use more clarification. We realized that a review of the evaluation life cycle stages and components were needed, as a number of grantees appeared confused over the distinctions. Additionally, it was apparent that most participants needed guidance on crafting appropriate evaluation questions, a critical part of any evaluation plan. We shared the following criteria for evaluation questions with the participants[4]:

Evaluation questions should be:

- ***Specific***. Does the question address a specific feature or intended outcome of the program?
- ***Measurable***. Will you be able to observe, gather data, and do the analysis needed to answer the question?
- ***Realistic***. Is the question realistic based on project lifecycle, timeline and experience of evaluator?
- ***Connected to the goal (relevant).*** Is the question clearly connected to the goal(s) set forth for the project?

A related activity was included with this discussion. Following the presentation and activity, participants had the opportunity to ask clarifying questions about their written feedback. Finally, the evaluation report template was reviewed and participants were offered an opportunity for feedback on draft reports if they were shared with us prior to the May meeting. All but one

[4] These criteria were adapted from suggested evaluation question guidance included in Wenzel and Brill (1998). Similar criteria are reported in Weiss (1972).

grantee submitted draft reports in time for the meeting. A feedback form was also used to help provide consistent feedback to the grantees on their draft evaluation reports.

During the May meeting, we shared reactions on the draft evaluation reports submitted for feedback. The feedback forms were distributed, and attendees were offered the opportunity to clarify questions or comments. The focus then turned to gathering reactions from the participants on the pilot project. This task included (a) questions on evaluation challenges faced and how they were addressed; (b) what participants would do differently in the future; (c) reactions to the evaluation guidelines, evaluation tools (logic modeling/strategy maps, evaluation plan and report template); and (d) overall suggestions to make the evaluation planning and reporting process more user friendly. Participants were also able to ask the BSF leader about the 2008–2009 application process.

Methods

The study was designed to capture multiple sources of qualitative data. For a large part of the pilot study, we acted as participant observers facilitating professional development sessions, probing for understanding of evaluation components and processes during meetings, and giving verbal and written feedback to participants on their evaluation efforts. During these sessions, we took note of participant reactions. In most sessions, two evaluation team members participated, allowing one to take notes while the other presented or facilitated group discussions. We also collected documents produced through the project, including original grant proposals and mid-year reports submitted prior to the pilot study, and evaluation plans and final reports using templates designed for the study.

We also conducted interviews with the director of external affairs and a sample of participants using a semi-structured interview protocol. The interview protocol is included in Appendix G. In addition to the BSF leader, a stratified random sample of six BSF project participants was selected for interviews to learn about their experiences and obtain feedback on the evaluation guidelines and tools. Two grantees were sampled from each of the three professional development groups: literacy, curriculum planning, and technology. (The two other project participants were assigned to the curriculum planning group). The interviews were held in the late spring and early summer of 2008, after the participants had completed their final reports. We made detailed notes as interviews were conducted and we used audio recordings to check for accuracy.

Data Analysis

Grant proposals, plans, and reports were analyzed using the QSR NVivo7 qualitative program. A coding structure was applied for data reduction and comparisons across projects on common elements (e.g., initial evaluation plans, evaluation questions). Coding also included elements from the approach applied during the first-year evaluation of the Trust Education Initiative. These codes were developed to categorize features of these documents according to their lifecycle on the grant continuum (e.g., planning/prototype, demonstration), and relevant evaluation features (e.g., context, input, process, impact, effectiveness). Additional codes were added during analysis to examine features of evaluation plans, report structures, and interview responses on their reactions to the professional development sessions, process and use of evaluation activities (including lessons learned) and documents. Written feedback provided to the grantees on their evaluation plans and reports was also coded to explore common problems as well as promising practices. Two members of the evaluation team coded these materials. Sample documents were cross-coded and compared to ensure consistency in coding interpretation and use.

Potential Limitations

This type of pilot study does have some limitations that should be noted. As participant observers, we were involved with both providing professional development to the grantees on the use of the evaluation tools and on evaluation concepts in general, while also testing the potential merits of the tools developed for the Trust. In order to check for our potential bias, we sought periodic feedback from the BSF director on the project to get her impressions of how the project participants were reacting to and using the tools.

The professional development included written and verbal feedback to the grant participants. A number of the participants noted that our feedback was helpful to them as they developed, executed, and reported on their evaluation activities. Therefore, the results from the pilot test of how "effective" the tools were need to be considered in the context of how the tools were shared with the participants, the feedback provided, and potential changes that resulted. We were careful to provide feedback that was insightful and not directive about potential changes to evaluation plans, activities, and reports. Nonetheless, it is possible that without some level of feedback or professional development before the new tools are widely used by the Trust or participating grantees, consistency of use and value to grantees may differ. The pilot study was not designed to establish causation, but instead to learn how the tools could be used, help grantees learn from their experiences, and get their reactions on feasibility and utility.

Findings

Findings from the pilot study were drawn from the document analysis and interviews. We reviewed project proposals, evaluation plans, mid-year and final reports, meeting interactions, and final interviews at the end of the pilot study with the six randomly selected project leaders and director of external affairs. Below we discuss the results as they pertain to each of our two main research questions.

1. How useful are the grant evaluation guidelines and related tools in helping grantees develop, monitor, and report on grant activities?

2. What are the implications for applying the guidelines and tools across the education program?

Usefulness of Evaluation Guidelines and Tools: Grantee Evaluations

Our analysis revealed a number of findings about the usefulness of the guidelines, tools, and features of the pilot study that included the professional development sessions. Main findings related to the following areas: (a) achieving more focused evaluations; (b) increasing accountability for results; (c) managing time and resources constraints; (d) improving the utility of the evaluation tools; (e) dealing with challenges with measures of student achievement and instructional change; (f) increasing attention to data collection methods and analysis; and (g) using evaluation components from the guidelines document.

Achieving more focused evaluations. The grantees prepared more focused evaluations and evaluation questions. As described below, by concentrating more on evaluations and using the guidelines project tools, grantees provided more detail on the aims and actions of their programs, and were able to make adjustments to their activities in response.

Each respondent reported that the evaluation process during the pilot project year was a substantial departure from the previous year. A typical comment was that: “We focused our questions much better than we did in the past. Before we really focused on telling a nice, feel good kind of story. We are still going to tell that story but with much more specific information” (Respondent 1). We found that participants fine tuned their evaluations throughout the guidelines pilot project. The six participants we interviewed all stated that this process led them to focus their evaluation plans and reports more efficiently.

The area of the evaluation process that experienced the most significant shift was the development and application of evaluation questions. Earlier grant procedures did not include evaluation questions specifically and most respondents did not report their evaluation questions. Respondent 3 explained that past plans addressed broad goals and aims, but did not address specific questions. Every grantee we interviewed mentioned the improvement in and importance of more focused evaluation questions. Respondent 2 explicitly stated that the most important element he took from [the evaluation team] was the idea that they needed to focus on questions they wanted to have answers for by the end of the year. Finally, Respondent 4 stated that “this process helped to clearly write-up evaluation questions.”

The evaluation questions did more than solely focus the evaluations; they also helped focus program activities. In the past, the lack of clear evaluation procedures and questions at times led to unspecific project goals. Respondent 6 commented that the guidelines project forced him to ask the most basic questions about why he was doing this project in the first place. He recalled when [an evaluation team member] looked at the first draft of his evaluation plan and asked him why he was using this particular solution to meet the school’s need and commented about how helpful it was to revisit that fundamental question. Similarly, Respondent 5 explained: “[We] didn’t have goals and objectives clearly labeled, it wasn’t specific. What [was] set up [in terms of the program itself] was good” but there were vague and confusing large paragraphs of text to describe program or evaluation aims rather than focused questions. The following section from the initial goal statement of this respondent’s proposal on their planned use of smart board technology illustrates this point:

> A current daily lesson includes opening a textbook, reading, assigning written exercises, and/or readings. Some lecturing and drilling on the parts by pointing at the text would occur. With interactive technology the lesson would have the skeleton on the screen, a student would leave his/her desk and manually drag labels to the matching parts of the skeleton. Or a student could print or write in bright colors at thicknesses appropriate for lower grades (and it will be converted to typewritten text with a click.) Then further

> interactive pages would be ready at a click to further teach the lesson with photos and manipulatives.

This paragraph gives detail on the use of the interactive technology, but does not outline the main goals of the project in a clear way. While participating in the pilot project and applying the evaluation tools, this grantees' project goal section turned into a primary goal, with sub-goals that were clear and measurable. Respondent 5's rewritten goals read:

> To provide teachers with the in-service and support needed to help them become proficient with the use of the [specific technology].
> - Teachers will use technology at a higher rate
> - The presence and use of equipment will positively change teacher's instructional practices
> - The rate and quality of student learning will improve, also limiting the amount of behavior disruption in class due to the increased level of classroom interaction
> - Teacher lessons will become more dynamic and also be more tailored to the individual student learner, opposed to a one-size-fits-all approach
> - Students will be encouraged to participate in higher order and critical thinking during interactive lessons

Respondent 5 was able to shift his project's goals through the application of basic criteria on constructing evaluation questions that was shared during one of the professional development sessions, and reinforced through written feedback.

By focusing on exactly what they would measure in terms of program effectiveness, Respondent 5 also focused on what they were aiming to achieve. Considering the guidelines model, the project's new question related to process, impact, and effectiveness evaluation questions, as seen in the following new evaluation questions:

1. Did teachers attend training sessions?

2. How often did teachers and students use boards?

3. Was the technology integrated into teachers' lesson plans?

4. Did student interest in learning increase?

Respondent 1 also stated that he was able to focus his program's goals by creating more specific evaluation questions. In his original project proposal, he laid out the project need and then launched into a description of the program elements, the implementation plan and timeline, and an evaluation/assessment plan. His original proposal never specifically stated the goals of the project or how the directors would know if they reached their goals. Their original evaluation plan included a vague and broad set of intents:

> [The project] will assess the impact of the [activities] by examining teacher retention, the quality of their instruction, and student growth as measured by academic performance. [We] will begin project implementation by administering a series of pre-surveys to both

> teachers and students. The surveys will contain questions and indicators unique to both teachers and children, gaining an initial snapshot of both educator and student attitudes and abilities (e.g. teacher satisfaction, students' active engagement in the learning process). [We] will administer post-surveys at the conclusion of the school year in late spring of 2007 (or more appropriate timeframe) to measure the impact of a school's respective professional development plan under the project.
>
> By achieving these outcomes, [we] and Big Shoulders will improve school and school community cultures to encourage greater teacher collaboration and promote an environment rich with relevant professional development opportunities. As a result, [we] and Big Shoulders will empower teachers to gain the knowledge and skills they need to become better educators, whether from their direct peers or experts external to their school. Over time, teacher retention will improve, quality instruction will increase, and students will begin to reach new levels of academic excellence in their studies.

As Respondent 1 explained, the guidelines pilot project fundamentally shifted this grantee's evaluation process. "Number 1: Fine-tuning really what you're looking for. This has been a big help for me. Rather than having one question that is huge and vague, it is probably better to have four really clean questions. I really had that 'a-ha' moment at the last session … fine-tuning, fine-tuning, fine-tuning. I made a big leap writing the report. Rather than trying to answer this really big vague question, I needed focused questions to measure." He created four clear evaluation questions to answer:

1. What value did teachers see in the professional development training provided by [the project]?

2. Are teachers using the strategies learned during the professional development training? How often?

3. Are the teachers seeing positive impact on student learning as a result of using the strategies learned? Where have they seen these results…grades, testing, etc…?

4. How has the professional development you have participated in this year affected your view of professional development training?

Similarly, Respondent 6 used the evaluation report template as a learning tool to reflect on his evaluation process and project. Under each evaluation question, he articulated his own learning process (in italics).

1. Do teacher lesson plans reflect an increase in competency? (More creative, evidence of co-planning and collaboration)

> *We discovered this question was vague and did not address the larger issue of student learning. Our methods for collecting data were not very sophisticated as we merely reviewed lesson plans on a weekly basis and determined (subjectively) if they were creative and/or reflected collaboration. This in no way indicated if the teachers were actually delivering creative, collaborative lessons to students. Further, we had no measures in place to determine the effectiveness of the lessons. This was a big red flag*

and a clear lesson learned. We definitely have to develop more focused questions and better measuring techniques.

2. Are teachers sharing knowledge with one another? (Dynamic meetings, evidence of sharing)

As with our competency question, our sharing question was too vague and difficult to collect measurable, valuable data. We did observe an increase in dialogue and interaction between teachers, but no indication that true collaboration was occurring. Again, collaboration was difficult to measure and we had nothing in place to measure how it impacted student learning. More thought has to go into the questions we ask; in particular, will the answers allow us to measure improved student performance?

This project has proved to be more of a learning experience than we might have imagined. And in many ways, it was a bit ambitious in terms of what we hoped to accomplish. Some aspects of our goals were met (acquiring and using Pasco equipment, improving teacher lesson plans, professional development using smartboards), but much of our goals have not been met. We learned that our questions need to be more focused and centered around student learning. We also have to develop better data collection and analysis processes. In terms of the scope of the project, we should have built in more time for professional development and training, where teachers become stakeholders and committed participants of the processes. In addition, we discovered that the scope of the changes we tried to implement, especially given the change in culture required, was more than we could accomplish in the course of one school year.

Finally, Respondent 4 similarly expressed the notion that the pilot process helped focus her project in addition to her evaluation:

> I realized my goals were just too much and too all over the place. It made me synthesize my questions and goals … It was excellent when you get down to it. Have a goal, what is the question you are going to ask, is it measurable, get the data and report on it.

While the evaluation questions were an area of growth throughout the grantees' evaluation plans and reports, there was still room for improvement. For example, as we see below, Respondent 2 included context and input evaluation questions but did not include any process, impact, or effectiveness questions.

- How will we assess effectiveness of teacher implementation of literacy instructional strategies in the classroom?
- How will we assess whether new tools/policies promote student progress toward improved literacy?
- How will we assess student progress toward specific literacy benchmarks (as measured by summative assessments indicated above)?
- How will teachers utilize collected data to inform future instructional planning?

- What metrics in addition to test scores can we use to define "success" in the context of this initiative?

These questions all look at how they plan to collect, measure, or use data but do not address how they will measure their project goals. Because this project was in the demonstration stage, they also needed to ask these types of impact or effectiveness questions. Though the professional development sessions, feedback forms, and evaluation tools engaged grantees in more concentrated work around evaluation in general, and evaluation questions specifically, as we see here grantees still have areas for growth.

Increasing accountability for results. The grantees felt more accountable in terms of engaging in and completing their programs and evaluating the results. We found that evaluation was not a central component of the grant process before the spring of 2007. By participating in the guidelines project and recognizing BSF and the Trust's support of the project, the participants' awareness of the importance of evaluation increased. Previously, evaluation was seen as a formal requirement to complete at the end of the grant period. But by participating in the professional development sessions, using the new evaluation templates, and receiving feedback, grantees thought more about evaluation and about their accountability for use of the grant funding.

An exchange with one participant illustrated these changes. Before, he stated that evaluation "… was very pro forma," with a mid-year report to check on progress and an end-of-year report. But, "Most of the data was qualitative or anecdotal in nature… it was more, 'this is how we used your money this year, and this is how we think things turned out as a result'" (Respondent 2). In contrast, he stated that this year forced his group to think much more critically about how they were using the resources and measuring results: "… for the first time we were asked to demonstrate 'what did success look like?' We really had to have goals we were specific about, questions we were specific about answering, then actual data we would present." He attributed this change to "…the introduction of [the evaluation team] into the process."

The grantees may have been previously exposed to evaluation through course work or experience, but the guidelines pilot project pushed them to be more explicit in measuring and documenting work on their projects. Respondent 4 explained, "I have learned a lot about evaluation in the past. I know about evaluation, have studied it." In fact, this grantee had written a master's thesis which included a large section on evaluation. But she explained that the guidelines pilot project "simplified everything and made it better." It connected what she already knew to this project and changed the way she wrote about her work. Respondent 5 concurred: "The narrative is useful but it isn't the whole thing and has to be backed up." The notion that their reports needed "to be backed up" with evidence grew during the pilot study.

Another respondent stated a similar notion. He had been exposed to evaluation processes as a consultant for a number of years, with heavy focus on continuous improvement. Though he stated that the tools the pilot study provided were similar to other quality tools he had been exposed to in the past, he found the tools very useful, because they forced him to continue to ask himself "why are you doing that?" As he stated, "you can always come up with things to do, but you need to ask why do you do it? If you have a sink overflowing in one room and then decide 'I'm going to vacuum, cut grass'… those are all good things, but how does it help with the sink?" (Respondent 6). The evaluation process forced him to think closely about the connection between the needs at his school and the goals and activities of the project the grant funded. Again, we see in his response the notion that what changed throughout this process was the accountability now central to the work.

Managing time and resource constraints. Respondents found that time and resources were the biggest obstacles to doing quality, in-depth evaluation work. When responding to a question about the biggest challenge, Respondent 5 wrote, "Time, because it was myself, gathering information from other schools … It's a long process. Having a development person would be important. Once I got into it, it pushed along but it took a lot of time." Similarly, Respondent 3 wrote: "Time is the biggest challenge especially when you are doing the kind of evaluation we are with videotape … obviously if this was the only thing on our plate … we could organize it well … it's not and so sometimes we scramble to make sure that we're getting the level of documentation that we really want." Resource constraints also challenged grantees' evaluation processes. Additional funding was not provided for evaluation purposes and often the school principal or project leader added evaluation and reporting tasks to multiple other responsibilities.

Turnover of teachers and school leadership is also an on-going issue. One project lost its project coordinator mid-way through the school year. This person was the lead evaluator for the project that had three schools. Therefore, Respondent 5 had to step in to complete the evaluation when she had not budgeted time for the project in her schedule. In addition, she did not have adequate background information from the initial project proposal on the evaluation plan. She also had to deal with resource challenges. Of the12 teachers who were involved in this technology initiative, 7 left the school. In addition, those teachers who were trained in technology during the prior year left their schools. As a multi-year project, this caused a setback because a number of teachers had to be trained in the fundamentals of the technology.

Improving the utility of the evaluation tools. The pilot project evaluation resources were helpful but could be improved in terms of format and practicality. In addition to the evaluation guidelines (based on the Lifecycle/CIPP format), the tools included the evaluation plan template, the evaluation report template, and the strategy map template. Overall, we found that while the plan and report templates were useful and important in helping the grantees focus their evaluations, the evaluation guidelines confused some grantees. Concerns were also expressed that the total amount of paper (the guidelines are 15 pages) was somewhat intimidating. In particular, the lifecycle determination was difficult for many participants. We found that we had to return to explain the lifecycle concept in feedback and during the second professional development session.

As Respondent 5 reported, "I was more than a little confused in terms of what lifecycle we're in. There was a lot of crossing between things – there was a little of each. [The evaluation team member] clarified it and I wrote down what he said." Similarly, Respondent 4 reported that:

> The bundling of papers was overwhelming to digest. Cut down the theory…. Be very specific and get to an example… If you had time to sit down and really study it that's great. But I had more time than most and I don't think those people are really going to do that. Tell me, what do I need to do?

Similarly, Respondent 7 noted her confusion about which form to use for the final evaluation, the report, or the plan template. Overall, this respondent felt that there were too many different forms and was not clear about the specific use and purpose of each form. The amount of paperwork and the organization of the paperwork did not make these tools easy for her to use. Another respondent mentioned that, "It took three readings to get the guidelines document, and I have done evaluation before. Repeating the components was confusing. It would be good to have it broken down by evaluation pieces" (Respondent 8). This person went on to say, "There's good stuff in guidelines, but they can be scared about the model or that there are repeat sections."

Other respondents saw the entire packet as a good resource. Respondent 3 reported that he used the guidelines document when there was a particular part of the template about which he wanted to make sure he was on the same wavelength as the evaluation team and BSF. He didn't use it in entirety, but dug deeply when he had a question. He too found the lifecycle determination difficult because he saw the demarcations as somewhat fluid.

Looking back over our feedback forms, we found the most confusion with the lifecycle determinations. As grantees wrote their evaluation plans, five grantees originally misidentified their lifecycle designation. It was difficult, though, to capture each grantee in one lifecycle and many grantees ended up marking two different stages, typically planning/prototype and demonstration. The primary difference between these two is the addition of effectiveness evaluation, though a planning/prototype evaluation would place more emphasize on context and input evaluation components. Respondents appeared to have resolved their lifecycle determinations by the evaluation report, but it isn't clear whether the grantees really understood the stages or simply adopted our feedback.

Dealing with challenges with measures of student achievement and instructional change. Measuring student achievement and instructional change were two challenges that were present throughout the evaluation pilot study. It was clear through our review of the initial evaluation plans, and during discussions with the participants, that it was expected each would try to link their professional development activities to some measures of student academic improvement. In addition, many also planned on measuring change in teacher instruction. However, there were often missing links between activities and student achievement. In some cases, adding student achievement measures to evaluation did not fit the scope of the project. For example, some projects were more focused on planning activities (which should not include effectiveness measures such as student achievement gains) because the end product is a plan for action. Others did not have the resources or capacity to interpret student achievement measures and link those to professional development activities.

Respondent 5 expressed her difficulty in using achievement scores to measure success:

> I still have difficulty with using test scores, I have a problem with that because with any program we put in, there are so many variables. Even though the test scores went up, I can not say it was because of the [program]. I would really like to see evaluation getting away from test scores, because it is impossible to isolate and be valid. Originally, [we] put in we would show that test scores went up. Tracking this would take at least 3 years, plus we get about 30% turnover from year to year.

During the professional development sessions, participants were encouraged to expand their evaluations to include process and impact features, not just an outcome related to student achievement (as many originally had). As noted, student achievement measures were not appropriate in each case. In addition, the grantees had not considered how they can measure links between their activities and ultimately student achievement. These links may include, for example, whether teachers are attending training (process focus), whether and how they are applying the training in their instruction (impact and effectiveness focus), and whether their instruction is changing as a result (effectiveness focus). Addressing these evaluation features would set them on the path to linking activities ultimately to student achievement. But given time constraints, we were not able to work with the projects on developing models to provide answers to how student achievement may be affected.

Increasing limited data collection and analysis. While we saw considerable growth in the development and follow-through on evaluation questions, we did not see as much growth in the data collection methods and even less in the data analysis. The past evaluation process did not address specific questions and tie them to specific data collection and data collection methods, according to Respondent 3. He continued to explain that in the past the focus was on writing the initial proposal and that he would adhere closely in terms of what the plan was, but was more skeletal about activities and did not address how questions would be answered by specific measures. Similarly, Respondent 6 wrote that in past plans the data collection methods and analysis section was easy to brush by without giving careful thought, but he stated that if you give some thought to collection methods in light of the other sections, you will be able to answer those questions more honestly.

Though we did push for closer links to the data collection methods, the tools were not specific enough in terms of data analysis, nor did we have enough time to delve into the topic during professional development. The evaluation report feedback forms frequently indicated that description of the data analysis was missing.

Usefulness of Evaluation Guidelines and Tools: Synthesis of Results

Analysis of the sample of six grantees indicated that, with few exceptions, each included evaluation components appropriate for their stage in the grant lifecycle.

A central feature of our evaluation tools is the lifecycle determination and corresponding components from the CIPP evaluation model. As we noted, grantees struggled with the concept of lifecycles and the placement of their program into specific lifecycle stages. After spending time on these issues in professional development sessions and on our feedback forms, grantees were able to correctly identify their lifecycle stage by the end of the year. Another analysis we completed was to determine how closely our six grantees stayed to reporting on the correct components that matched their lifecycle stage. We mostly analyzed their evaluation final reports to look for this evidence.

The context component was included in all six grantees' evaluations. One section of the report explicitly included background and context and each grantee reported background information and program goals. Four of the six grantees included information about input evaluation. One grantees' final reported included no information about how the specific strategies they decided to employ best met the goals and budget of the grant while one other grantee (Respondent 1) noted that they employed "responsive programming … designed to be unique to the needs of each school and its student" but did not provide any further details. The other four grantees in our sample explicitly stated why this program best met their goals and needs. For example, Respondent 6 considered two questions: How will your planned activities affect progress in achieving program goals and what are the main links between the activities and the goals that will lead to success in meeting your objective. Respondent 5 wrote that the technology system they chose "has been shown to be a most effective and economical way for total class involvement, engaged learning, with immediate, formative assessment feedback."

Process evaluation was a strong component for these grantees, with five of the six grantees discussing issues of monitoring, documenting, and assessing program activities in their final reports. For example, Respondent 5 reported that "teachers sign in at each training session and sheets are collected. Teachers sign schedules at each school to use Promethean board. Times and dates are indicated on schedules. Collection and analysis of these schedules indicated how often the boards were used and by what teachers." Respondent 3 was able to integrate process evaluation and their program goals, as they explained, "the purpose of the faculty portfolio was to act as both a framework for developing shared expectations around classroom practice (a program goal) as well as housing documentation of goal development and completion." Interestingly, Respondent 6 reflected upon the fact that they did not have enough process evaluation in place; this fact hindered their ability to report on impact or effectiveness. They explained,

> Our methods for collecting data were not very sophisticated as we merely reviewed lesson plans on a weekly basis and determined (subjectively) if they were creative and/or reflected collaboration. This in no way indicated if the teachers were actually delivering creative, collaborative lessons to students. Further, we had no measures in place to determine the effectiveness of the lessons. This was a big red flag and a clear lesson learned. We definitely have to develop more focused questions and better measuring techniques.

One of our goals was to help grantees distinguish between impact and effectiveness evaluation activities. However, effectiveness was included more often in the reports than impact. Five out of the six grantees asked and answered effectiveness questions while only three asked and answered impact questions. One grantee raised both impact and effectiveness questions in his program goal section of the report but did not report on these issues. Another grantee mentioned in his interview that impact questions were important and admitted that he often overlooked these, and he did not report them in his evaluation report. Respondent 5 noted how many teachers attended training sessions and how often teachers and students used the boards. The lack of attention to the impact questions limited the ability of BSF to aggregate results across grants in terms of teachers, students, and principals who were impacted by the funded professional development activities.

We provided a number of example questions about effectiveness. A few respondents incorporated questions about how effective the program activities were in meeting goals in their reports. For example, from Respondent 1: *Are the teachers seeing positive impact on student learning as a result of using the strategies learned?;* from Respondent 5: *Were teacher's* *instructional practices positively changed? Were lessons more dynamic?*; and from Respondent 4: *What percentage of teachers being trained have changed instructional styles?*

We believed the lifecycle stages and corresponding components were an important guide for evaluation. While many grantees understood and incorporated them into their evaluations, there is clearly more work to do to cement these concepts.

Another primary goal of the project was to demonstrate how the guidelines and report templates could be used to better judge the progress and outcomes of the grants funded. In this respect, we would be providing evidence that would help the Trust examine progress across grants to determine success of its funding strategy and help with decisions about future grants. BSF changed how they monitored and reported their evaluation results and experienced some improvements and challenges. It was apparent that synthesizing results using the guidelines and tools remains a challenge. In this section, we briefly discuss the results of participating in the pilot study and implications for applying the tools across projects.

BSF used a revised evaluation plan developed with the overall goal of helping BSF schools, principals, and teachers become more effective educators. The subject of the evaluation efforts was on the evaluation pilot study activities and related professional development sessions. This would allow them to (a) see how effective the process was working for their grantees, (b) help assess a goal of developing capacity within schools to evaluate their own programs, and (c) use the information to share best evaluation practices and school improvement strategies. In effect, collecting and analyzing the evaluation plans, reports, and future proposals would help BSF assess the extent to which their overall professional development program has been successful and uncover needs for additional support. The expected outcomes of BSF also changed. For example, rather than focusing on an outcome related to improving student motivation, as in the original BSF proposal (shown on page 17), the first outcome listed in the plan was to: “Involve teachers and principals in designing and deciding upon appropriate professional development programs, goals and evaluation plans.”

BSF’s final report to the Trust was similar to those submitted in recent years, with an overall report on BSF as an organization, a brief report on the evaluation pilot study from their perspective, then a project-by-project review of grant activities and outcomes. The report included aggregate results with descriptives on the amount of funding allocated, numbers of teachers, students and principals directly and indirectly impacted by the professional development funding (including trends on these measures over three years). Then, for each BSF project, a summary report was included on the project award amount, project goal, project implementation, project impact (number of schools, students, leaders), and project evaluation summary. The main difference between these reports and those of years past was the inclusion of a more specific evaluation summary and the full evaluation reports completed by individual projects in an appendix.

According to a BSF leader, the key difference between evaluation this year and in previous years was that this year was more process oriented. The schools had more technical assistance and steps to complete an evaluation than ever before. In the past, quality varied considerably and even though there was still variability in the quality of reports this year, it “was important to give schools tools to use for evaluation, the logic model is helpful, and the ‘how to’ is very important to our schools." Their educators have the desire, commitment, and dedication, but in the past evaluations frequently reported that: "kids are more excited, teachers are more excited, they got a lot out of it." This process held them to a higher standard in conducting and reporting their evaluations.

The primary concern for BSF in its report to the Trust was the difficulty it had in aggregating common indicators from the project reports. The schools still struggle with how to report student achievement changes linked to professional development activities. In addition, schools would have benefited if the final report template had included a specific prompt for numbers of teachers and students impacted and the basic demographics of the students. Despite the help of the reporting tools in getting more consistent reports, information was still hard to synthesize across all grants. Summing up the experience, the BSF leader explained that the professional development grant process is getting more competitive, which is good. Schools are becoming more focused, and others clearly need more support developing proposals. Continuing to focus on how they support evaluation will be an on-going goal.

When evaluators attempt to assess activities and progress for formative uses (to change and grow a program, for example), they encounter tension between this activity and the simultaneous need to obtain data from grantees that can be synthesized across grants. This issue was raised about the current guidelines and reporting tools during our interviews and as we analyzed evaluation plans and reports. This issue will also be apparent with the evaluation strategy as the Trust moves ahead with its new education program. For example, the original evaluation plan for BSF had broad indicators of success that each BSF project was intended to contribute, including (a) measures of teacher satisfaction, (b) teacher retention, (c) student motivation and engagement, and (d) student achievement. Many projects neglected to report how many teachers, students, and administrators were "reached" by their projects. While intended beneficiaries are called for in the evaluation plan, and are a key component of Process and Impact evaluation according to the guidelines, results on beneficiaries served are not prompted in the evaluation report. This should be specifically added. BSF also struggled with which measures of student achievement were appropriate to include, and how to attribute changes in achievement. These are key indicators that BSF was seeking to address, and issues that other funders want to see included in grant reports.

One way to address these challenges would be the development of overarching strategic plans for grant-making. BSF would benefit from developing logic models for their grant-making strategy. These models, combined with performance indicators, can then be used to structure evaluation strategies that align the Trust's funding within their organizational strategies to improve member schools. Doing so will help build a more focused trajectory for grant funding within the organization. In addition, the process could better align with and support evolving logic models under development by the Trust.

Summary of Findings and Recommendations

The evaluation guidelines pilot project has helped test the evaluation tools, and has also provided information about the type of supports many smaller grantees need to do quality evaluation work, both in terms of the way the tools are interpreted, and the guidance that can help produce results for the organization, its stakeholders and the funding organization. Key findings and recommendations from the pilot project include following:

1. In the past, most grantees did not focus on their evaluations until the end of the grant and saw the final evaluation primarily as a reporting requirement designed to show how funds were used and whether the activities were perceived as successful or not. This precluded using emerging findings from evaluation to refine activities during the grant year or in subsequent years. With little guidance, grantees submitted reports and evaluation plans of varying depth and quality. Our analysis comparing evaluation reports before and after the pilot indicates that grantees provided more specific evaluation questions, were more likely to justify findings with data, and appeared to learn from their evaluation activities. The evaluation plan and report frameworks were seen as useful and helped project leaders remain focused on their core activities for their grants.

   Recommendation: With minor modifications and guidance, the evaluation plan and report templates can be tested more widely.

2. Beyond the evaluation tools, grantees reported that they acquired useful knowledge from participating in the evaluation guidelines project. Several pointed out that the feedback provided was instrumental in helping them specify evaluation activities and questions that would be most informative and feasible for their projects. Others saw how the process could inform other projects in their organizations. Still others valued the timelines involved in the project to keep them on track.

   Recommendation: Ideally, continued professional development and guidance would help smaller grantees get the most out of their evaluations, and help establish more uniform evaluation reporting. This can include a general overview for new grantees to get them started on evaluation planning, with follow-up feedback as needed. At a minimum, user guides on the evaluation plan, report, and logic model should be developed and included with the Web-based grant submission process.

3. At first exposure, the evaluation guidelines document was confusing to most participants without explanations and opportunities to ask clarifying questions. Several grantees got stuck trying to determine the appropriate lifecycle and related evaluation components for their projects. The format for the guidelines was also seen as cumbersome. When their use as a reference was clarified, participants were more comfortable.

   Recommendation: The guidelines document should remain as a resource for grantees, especially if they are new to evaluation and confused about where to start. They should be formatted, however, for ease of reference using Web technology (point and click with pop-ups) that can logically lead a grantee from the lifecycle stage to the appropriate components when they are planning their evaluation. A Web-based format will address the intimidation some feel when viewing the longer document. For proposal submissions, grantees could indicate where they think they are on the grant life-cycle continuum, but program officers should verify and clarify the designation. Grantees need to know that the guidelines are a tool to help them get in the right “ballpark” for evaluating their programs, but they still need to develop evaluation questions to address the most important grant activities to assess, learn from, and inform funders and other stakeholders. The evaluation plan and report template will help them focus on the most important evaluation activities to address. Other resources, including the Program Evaluation Standards of the Joint Committee for Educational Evaluation (http://www.jcsee.org/program-evaluation-standards), should also be highlighted for grant recipients in their development and reporting of evaluations or work with external evaluators.

4. Although suggested as an optional tool, most who attempted to develop a strategy map/logic model saw that the process was helpful in reflecting on planned activities and formulating evaluation plans. Some were confused by the original version and suggestions were made to revise the template.

   Recommendation: The logic model format has been simplified. It can serve as a tool for all grantees to complete as they prepare their grant proposals, especially the initial evaluation plan. Examples of logic models from Trust-funded projects, and other widely available resources on logic model applications for evaluation, can illustrate use.

5. When evaluators attempt to assess activities and progress for formative uses (to change and grow a program, for example), they encounter tension between this activity and the simultaneous need to obtain data from grantees that can be synthesized across grants. This issue was raised about the current guidelines and reporting tools during our interviews and as we analyzed evaluation plans and reports. This issue will also be apparent with the evaluation strategy for the new Trust education program.

   Recommendation: Adaptations to the instruments are needed to identify common indicators all projects should report. In the case of BSF, this should include common metrics for numbers of students, teachers, and leaders impacted, and for student achievement. However, this data could also be collected by BSF. The challenge will be linking teachers and students to funded projects. That will require adoption of a data system that can warehouse school and classroom level data (if possible) by the Archdiocese and accessible by BSF.

6. Funded projects would benefit from developing logic models that “fit” within the strategic plan and logic models of the Trust, with performance indicators feeding into evaluation plans developed from the models. The program officers can work with the grantees to develop the models.

   Recommendation: Grant projects should develop overarching logic models for their grant-making strategy, with the participation of the program officers.

Evaluation is a challenging and evolving field, with a number of theoretical approaches that could guide activities. Foundations typically struggle with questions of how, when, and for what purposes to evaluate grant-making activities. Rather than embarking on strategic, on-going evaluation efforts, if they do evaluate their programs, many attempt to periodically assess the impact of an entire portfolio of grants, or selected high profile grants, and become frustrated when evaluations can not attribute causation. Frustrations may be compounded due to the difficulty in quantifying the impact of nonprofit, social service organizations, which often have the broad goal of improving people’s lives (Kramer et al., 2007). This project has demonstrated the design and use of one set of evaluation tools that can apply to grants of different size, foci, and stage in the project lifecycle. With refinement, the guidelines and tools could also be used by foundation staff to help assess portfolios of grants or initiatives. Additional testing of these tools and application by other organizations will help to further document their potential to meet evaluation challenges.

References

Behrens, T. R., & Kelly, T. (2008). Paying the piper: Foundation evaluation capacity calls the tune. *Nonprofits and Evaluation*, (119), 37–50.

Carman, J. G. & Fredericks, K. A. (2008). Nonprofits and evaluation: Empirical evidence from the field. *Nonprofits and Evaluation*, (119), 51–71.

Condliffe Lagemann, E., & de Forest, J. (2007). What might Andrew Carnegie want to tell Bill Gates? Reflections on the hundredth anniversary of the Carnegie Foundation for the Advancement of Teaching. In R. Bacchetti & T. Ehrlich (Eds.), *Reconnecting education & foundations: Turning good intentions into educational capital* (pp. 47–70). San Francisco, CA: Jossey-Bass.

Hendricks, M., Plantz, M. C., & Pritchard, K. J. (2008). Measuring outcomes of United Way-funded programs: Expectations and reality. *Nonprofits and Evaluation*, (119), 13–35.

Joint Committee on Standards for Educational Evaluation. (2010). *Program evaluation standards: Summary.* Retrieved from http://www.jcsee.org/program-evaluation-standards.

Kramer, M. R., & Bickel, W. E. (2004). Foundations and evaluation as uneasy partners in learning. In M. T. Braverman, N. A. Constantine, & J. K. Slater (Eds.), *Foundations and evaluation: Contexts and practices for effective philanthropy* (pp. 51–75). San Francisco, CA: Jossey-Bass.

Kramer, M., Graves, R., Hirschhorn, J., & Fiske, L. (2007). *From insight to action: New directions in foundation evaluation*. FSG Social Impact Advisors. Retrieved from http://www.fsg-impact.org/

Kronley, R. A. (2000). *Learning from each other: Questions and answers about the Clark, Panasonic and Rockefeller Foundations' long-term investment in systemic education reform*. Atlanta, GA: Author.

Raymond, Margaret (2001). Evaluating education philanthropy in action. In J. Barry, P. Frumkin, P. T. Hill, L. Lenkowsky, B. V. Manno, M. Raymond, . . . & J. Williams (Eds.), *Seven Studies in Education Philanthropy.* Retrieved from http://www.edexcellence.net/institute/publication/publication.cfm?id=318&pubsubid=867&doc=pdf .

Stufflebeam, D. L., Madus, G. F., & Kellaghan, T. (2000). *Evaluation models: Viewpoints on educational and human services evaluation* (2nd ed.). Boston, MA: Kluwer Academic Publishers.

W. K. Kellogg Foundation (1998). *W. K. Kellogg Foundation evaluation handbook.* Battle Creek, MI: W. K. Kellogg Foundation. Retrieved from http://www.wkkf.org/Pubs/Tools/Evaluation/Pub770.pdf

Weiss, C. H. (1972). Evaluation research: Methods of assessing program effectiveness. Englewood Cliffs, NJ: Prentice Hall.

Wenzel, S. A. & Brill, E. (1998). Mapping a path for evaluation: A planning guide (2nd ed.) Chicago, IL: Girl’s Best Friend Foundation.

Appendix A

Draft CCT Evaluation Guidelines

The Chicago Community Trust requires evaluations of the programs we fund. In order to be successful in evaluating your program and reap the benefits that evaluation can offer, it is important to understand what program evaluation is, what benefits evaluation can produce for your program or organization, and how and why we use evaluations at the Trust.

What is Program Evaluation?

Evaluation is defined in many ways and used in a variety of disciplines to understand, improve, and measure the quality, process, impact, and effectiveness of a program, project, or initiative. Put simply, program evaluation is the systematic investigation of the worth or merit of a program, project, or instructional material.[5] Program evaluation can serve both short and long-term projects and activities. The Trust's evaluation requirements are intended to be adaptable to your program, regardless of your project's size, duration or type.

What Value Does Program Evaluation Have for Your Project?

Evaluation activities can help identify a problem to be addressed, produce alternatives to addressing the problem, judge the merits of alternatives, study the results, and generate changes based on what has been learned.[6] Evaluation feedback can assist program leaders and staff in ensuring the delivery of quality services to identified beneficiaries. By incorporating quality evaluation into the purpose and activities of the program, identified stakeholders can assess the progress and impact of the program and use this valuable feedback to improve processes, redirect services or resources, plan for the future of the program, and foster overall organizational growth and development. It is important that programs identify, learn from, and communicate missteps as well as successes. The resulting information is not only useful for the program, but also for the Trust as it strives to help the grantee and other grantees in future activities.

Why is Program Evaluation important to The Chicago Community Trust?

Just as evaluation benefits grant program leaders and grant program beneficiaries, it is also an important component of building knowledge, funding decisions and strategic planning at The Chicago Community Trust. Through evaluation of the programs we fund, the Trust aims to gain a better understanding of how the programs we support are working and impacting the community. Learning what programs and approaches are effective at addressing challenges in the Chicago community helps us direct our efforts to promote these successes, learn from the challenges, provide information to our board, and evaluate applications for funding. The following program evaluation guidelines are designed to make transparent our evaluation

---

[5] The Program Evaluation Standards, 2nd Edition, Joint Committee on Standards for Program Evaluation. Thousand Oaks, CA: Sage

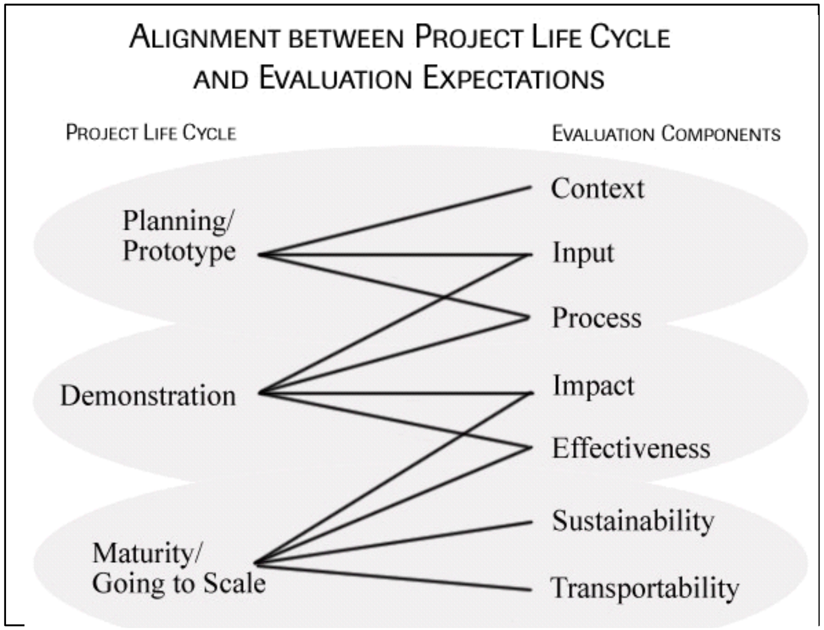
[6] Shadish, W.R., Cook, T.D. & Leviton, L.C. (1991). Foundations of Program Evaluation: Theories of Practice. Newbury Park, CA: Sage.

expectations, ensure some consistency across evaluations of CCT projects, and generate an evaluation process that is feasible and useful.

Evaluation Requirements*

What are the evaluation requirements?
Our evaluation expectations apply to evaluations that are either internally conducted by the grantee or external to the grantee. The evaluation expectations are guided by the program's stage in the Project Life Cycle. As shown by the figure below, a program can be considered in stages of Planning/Prototype, Demonstration, or Maturity/Going to Scale.

Not every program the Trust funds will go through each life cycle stage. However, the life cycle stage of a program is central to determining what evaluation questions are appropriate to ask when assessing the program. The first step in determining an appropriate evaluation strategy is to think about program goals and how program activities are expected to impact progress in achieving these outcomes. Keeping these points in mind, the following Life Cycle definitions and evaluation components will help in identifying an evaluation strategy.



Planning/Prototype: This category includes planning, research and development and/or initial implementation of a new program. Examples of prototypes include research for an intervention on school drop-out prevention or planning activities to develop and start a charter school. An evaluation of a planning/prototype program includes examining its context, inputs, and process. Programs in the planning/prototype phase are generally not ready to be evaluated for impact or effectiveness.

Demonstration: The demonstration stage covers programs that have been developed, but not yet thoroughly tested in the targeted site(s). An example of a demonstration project is a professional development program that was designed based on relevant research literature and practical experiences working with schools, but not yet applied in its entirety. Demonstration evaluation involves assessing inputs, process, impact, and effectiveness of the program, as well as its potential for continuation or expansion.

Maturity/Going to Scale: This phase applies to programs, initiatives, models, or specific features of each, that have been thoroughly tested in their demonstration stage and have reached maturity. It entails applying what is "known" to work about a project or program in a given setting (or settings), and applying the knowledge on a larger scale. These programs may be expanding to the whole organization or larger numbers of organizations and may include adaptations to fit the adopting organization(s). Programs in this stage may involve a substantial funding commitment and possibly include multiple funders. An example of such a project is a literacy program that has been implemented and tested in a number of schools and is being fully implemented in one or more area units. Evaluation in this stage involves examining the program's impact, effectiveness, sustainability, and transportability.

Each of these life cycle stages is associated with applicable evaluation components. Grantees will work with the Trust to determine their programs' lifecycle and which evaluation components are appropriate.

*Adapted from the CIPP Evaluation Model Checklist, Stufflebeam (2002), http://www.wmich.edu/evalctr/checklists/cippchecklist.htm

Planning/Prototype

This category includes planning, research and development and/or initial implementation of a new program. Examples of prototypes include research for an intervention on school drop-out prevention or planning activities to develop and start a charter school. An evaluation of a planning/prototype program includes examining its context, inputs, and process. Programs in the planning/prototype phase are generally not ready to be evaluated for impact or effectiveness.

Context

Purpose: For a program in its initial stages, examining the context of the program within a defined environment involves assessing the needs and assets of program beneficiaries and participants, as well as identifying problems (e.g. readiness for change, limited capacity) the program will need to solve to succeed. Results can be used throughout the program and at its conclusion to assess effectiveness and significance in meeting assessed needs.

<table>
<tr><th>Evaluation Activities</th><th>Activity Questions</th><th>Use</th></tr>
<tr><td>1. Compile and assess background information, especially on the intended beneficiaries’ needs and assets and community environment.
→</td><td>-What are the needs and assets of the program beneficiaries based on review of background information?
-What are the relevant characteristics of the community environment?
-What background information/data was used to determine needs and assets?</td><td rowspan="3"><u>Use Context findings to:</u>
- Select the intended beneficiaries
- Review appropriateness of program goals
- Ensure that the program will build on community and other assets
- Monitor program effectiveness and significance in assessed needs</td></tr>
<tr><td>2. Interview program leaders and other stakeholders to discuss their perspectives of beneficiary needs, assets and potential problems with the program.
→</td><td>-Who was interviewed and why? What needs did they identify?
-How do they relate to needs indicated from the review of background information?
- What problems do they predict in implementing the program?</td></tr>
<tr><td>3. Assess program goals
→</td><td>-Are program goals applicable to the assessed needs and assets of identified beneficiaries and the given environment?</td></tr>
</table>

Inputs

Purpose: Input evaluation assesses competing strategies considered for the program as well as the work plans and budgets of the selected strategy/approach.

Evaluation Activities	***Activity Questions***	***Use***
1. Identify and investigate existing programs that could serve as a model for the proposed program. →	-What existing programs were examined? -Why were these programs considered as models for the proposed program? -How were the comparison programs implemented? -What results did the comparison programs achieve? -Is the proposed program a new approach (i.e. no model for it exists)?	<u>Use Input findings to:</u> - Refine proposed strategy to be scientifically, economically, socially, politically, and technologically defensible - Ensure feasibility of meeting the assessed needs of the beneficiaries - Ensure financial feasibility - Ensure that the program is
2. Assess proposed strategy for responsiveness to assessed needs. →	-What is the proposed program model or strategy that was selected? -What is the program's theory of action or logic model? -How does the proposed program model or strategy address the needs, assets and problems identified in the Context evaluation?	
3. Assess proposed strategy for rigor. →	-How does the research literature inform the structure, plan, and goals of the program? -How does the proposed strategy compare to alternative strategies and related programs?	

4. Assess program feasibility. →	-Is the proposed budget sufficient to fund the work of the program? -Is the work plan and schedule for the proposed program sufficient to serve the program's goals, feasible with available resources, and politically viable? - Does the choice of sites/participants fit within existing structures? - Does the program align with other existing programs and/or organizations' goals?	superior to its alternatives and uses a rigorous strategy

Process

Purpose: Evaluating the program's process involves monitoring, documenting and assessing program activities as the program is taking place.

Evaluation Activities	***Activity Questions***	***Use***
1. Monitor, observe, document progress of program implementation. →	-What observations were made about how the program is being implemented? -How often were program activities monitored/observed? -What methods of monitoring/observation were used and why? -What progress has been made in program implementation? -What strategic or programmatic changes or adjustments have been made during program implementation? -What program challenges were encountered and how were they addressed?	<u>Use Process findings to:</u> - Record and report on progress - Reflect on evaluation activities - Monitor and improve implementation - Strengthen program design for possible
2. Interview beneficiaries, leaders, and staff to obtain their assessments of the program's progress. →	-What did leaders, beneficiaries, and program staff report about implementation progress? -Are these assessments aligned with the evaluation observations?	

3. Record events, problems, costs, and allocations. Provide copies of these records in the evaluation report as applicable. →	-What methods were used to maintain records on program events, problems, costs, and allocations? -What events, problems, costs, and allocations were found?	continuation -Disseminate results -Share knowledge -Create opportunities for feedback
4. Disseminate results to appropriate constituencies or stakeholders. →	-Is there a specific dissemination plan to share both interim and final evaluation findings? -How do program planners intend to use the findings? -Do program planners/staff dedicate time to reflect on evaluation findings? -What steps were taken to disseminate results? What stakeholders were identified to receive results?	

Demonstration

The demonstration stage covers programs that have been developed, but not yet thoroughly tested in the targeted site(s). An example of a demonstration project is a professional development program that was designed based on relevant research literature and practical experiences working with schools, but not yet applied in its entirety. Demonstration evaluation involves assessing inputs, process, impact, and effectiveness of the program, as well as its potential for continuation or expansion.

Inputs

Purpose: Input evaluation assesses competing strategies considered for the program as well as the work plans and budgets of the selected strategy/approach.

Evaluation Activities	***Activity Questions***	***Use***
1. Identify and investigate existing programs that could serve as a model for the proposed program. →	-What existing programs were examined? -Why were these programs considered as models for the proposed program? -How were the comparison programs implemented? -What results did the comparison programs achieve? -Is the proposed program a new approach (i.e. no model for it	<u>Use Input findings to:</u> - Refine proposed strategy to be scientifically, economically,

	exists)?	socially, politically, and technologically defensible - Ensure feasibility of meeting the assessed needs of the beneficiaries - Ensure financial feasibility - Ensure that the program is superior to its alternatives and uses a rigorous strategy
2. Assess proposed strategy for responsiveness to assessed needs. →	-What is the proposed program model or strategy that was selected? -What is the program's theory of action or logic model? -How does the proposed program model or strategy address the needs, assets and problems identified in the Context evaluation?	
3. Assess proposed strategy for rigor. →	-How does the research literature inform the structure, plan, and goals of the program? -How does the proposed strategy compare to alternative strategies and related programs?	
4. Assess program feasibility. →	-Is the proposed budget sufficient to fund the work of the program? -Is the work plan and schedule for the proposed program sufficient to serve the program's goals, feasible with available resources, and politically viable? - Does the choice of sites/participants fit within existing structures? - Does the program align with other existing programs and/or organizations' goals?	

Process

Purpose: Evaluating the program's process involves monitoring, documenting and assessing program activities as the program is taking place.

Evaluation Activities	***Activity Questions***	***Use***
1. Monitor, observe, document progress of program implementation. →	-What observations were made about how the program is being implemented? -How often were program activities monitored/observed? -What methods of monitoring/observation were used and why? -What progress has been made in program implementation? -What strategic or programmatic changes or adjustments have been made during program implementation? -What program challenges were encountered and how were they addressed?	<u>Use Process findings to:</u> - Record and report on progress - Reflect on evaluation activities - Monitor and improve implementation - Strengthen program design for possible continuation -Disseminate results -Share knowledge -Create opportunities for feedback
2. Interview beneficiaries, leaders, and staff to obtain their assessments of the program's progress. →	-What did leaders, beneficiaries, and program staff report about implementation progress? -Are these assessments aligned with the evaluation observations?	
3. Record events, problems, costs, and allocations. Provide copies of these records in the evaluation report as applicable. →	-What methods were used to maintain records on program events, problems, costs, and allocations? -What events, problems, costs, and allocations were found?	
4. Disseminate results to appropriate constituencies or stakeholders. →	-Is there a specific dissemination plan to share both interim and final evaluation findings? -How do program planners intend to use the findings? -Do program planners/staff dedicate time to reflect on evaluation findings? -What steps were taken to disseminate results? What stakeholders were identified to receive results?	

Impact

Purpose: Evaluating the program's impact involves assessing a program's reach to the *target* audience.

Evaluation Activities	***Activity Questions***	***Use***
1. Maintain a directory of persons/groups served, make notations on their needs, and record program services they received. →	Include in ongoing documentation system: -Who is being served by the program? -What are their needs that can be served by the program? -What services are they receiving or did they receive? -When and for how long did they receive these services?	<u>Use Impact findings to:</u> - Assess the extent to which the program is serving targeted persons/groups and addressing diagnosed community needs and goals - Redirect programming to intended beneficiaries as needed
2. Assess and make a judgment of the extent to which the served persons/groups are consistent with the program's intended beneficiaries. →	-Are the program beneficiaries the same as those intended to be served as outlined in the goals of the program (e.g. from the Planning/Prototype phase, see Context evaluation)? -Who are these individuals/groups and how do you know they are the intended beneficiaries?	
3. Interview area stakeholders to learn their perspectives on how the program is influencing the community. →	-What are the stakeholders' assessments of how the program is influencing the community? -Ask interviewees to provide specific examples of program influence.	
4. Determine the extent to which the program inappropriately provided services to a non-targeted group. →	-What persons/groups were served by the program who were not considered a targeted group at the outset/goal-setting phase of the program? -How do you know they are not the intended beneficiaries? -Why were these persons/groups provided services? - Were there any unintended services provided? If so, why?	

Effectiveness

Purpose: Effectiveness evaluation assesses the quality and significance of outcomes.

Evaluation Activities	***Activity Questions***	***Use***
1. Interview key stakeholders, such as community leaders, beneficiaries, program leaders and staff, to determine their assessments of the program's positive and negative outcomes. →	-What are the key stakeholders' assessments of the program's outcomes? -Are these assessments in line with the evaluator's findings?	Use Effectiveness findings to: - Gauge the program's positive and negative effects on beneficiaries - Gauge the program's positive and negative effects on the community - Judge important unexpected effects of the program - Examine whether program plans and activities need to be changed
2. Collect information from program documentation and participants to identify, confirm, and analyze the range, depth, quality, and significance of the program's effects on beneficiaries and potential for continuing the program. →	-Who are the beneficiaries? -What are/were the beneficiaries' needs? -What services did the beneficiaries receive from the program? -To what extent did the person/group benefit from the program? What were the intended/unintended and positive/negative impacts on the person/group due to the program? -What impacts/effects on the person/group were observed/recorded? -What are the appropriate measures of quality? -What are the methods used for measuring quality? -What steps have been taken toward assessing and/or ensuring that the program can continue?	

Maturity/Going to Scale

This phase applies to programs, initiatives, models, or specific features of each, that have been thoroughly tested in their demonstration stage and have reached maturity. It entails applying what is "known" to work about a project or program in a given setting (or settings), and applying the knowledge on a larger scale. These programs may be expanding to the whole organization or larger numbers of organizations and may include adaptations to fit the adopting organization(s). Programs in this stage may involve a substantial funding commitment and possibly include multiple funders. An example of such a project is a literacy program that has been implemented and tested in a number of schools and is being fully implemented in one or more area units. Evaluation in this stage involves examining the program's impact, effectiveness, sustainability, and transportability.

<u>Impact</u>

Purpose: Evaluating the program's impact involves assessing a program's reach to the *target* audience.

Evaluation Activities	***Activity Questions***	***Use***
1. Maintain a directory of persons/groups served, make notations on their needs, and record program services they received. →	Include in ongoing documentation system: -Who is being served by the program? -What are their needs that can be served by the program? -What services are they receiving/did they receive? -When and for how long did they receive these services?	<u>Use Impact findings to:</u> - Assess the extent to which the program is serving targeted persons/groups and addressing diagnosed community needs and goals - Redirect programming to intended beneficiaries as needed
2. Assess and make a judgment of the extent to which the served persons/groups are consistent with the program's intended beneficiaries. →	-Are the program beneficiaries the same as those intended to be served as outlined in the goals of the program (e.g. from the Prototype phase, see Context evaluation)? -Who are these individuals/groups and how do you know they are the intended beneficiaries?	
3. Interview area stakeholders to learn their perspectives on how the program is influencing the community. →	-What are the stakeholders' assessments of how the program is influencing the community? -Ask interviewees to provide specific examples of program influence.	
4. Determine the extent to which the program inappropriately provided services to a non-targeted group. →	-What persons/groups were served by the program who were not considered a targeted group at the outset/goal-setting phase of the program? -How do you know they are not the intended beneficiaries? -Why were these persons/groups provided services? - Were there any unintended services provided? If so, why?	

Effectiveness

Purpose: Effectiveness evaluation assesses the quality and significance of outcomes.

Evaluation Activities	***Activity Questions***	***Use***
1. Interview key stakeholders, such as community leaders, beneficiaries, program leaders and staff, to determine their assessments of the program's positive and negative outcomes. →	-What are the key stakeholders' assessments of the program's outcomes? -Are these assessments in line with the evaluator's findings?	Use Effectiveness findings to: - Gauge the program's positive and negative effects on beneficiaries - Gauge the program's positive and negative effects on the community - Judge important unexpected effects of the program - Examine whether program plans and activities need to be changed
2. Collect information from program documentation and participants to identify, confirm, and analyze the range, depth, quality, and significance of the program's effects on beneficiaries. →	-Who are the beneficiaries? -What are/were the beneficiaries' needs? -What services did the beneficiaries receive from the program? -To what extent did the person/group benefit from the program? What were the intended/unintended and positive/negative impacts on the person/group due to the program? -What impacts/effects on the person/group were observed/recorded? -What are the appropriate measures of quality? -What are the methods used for measuring quality? -What steps have been taken toward assessing and/or ensuring that the program can continue?	

Sustainability

Purpose: Sustainability evaluation assesses the extent to which a program's contributions are successfully institutionalized and continued over time.

Evaluation Activities	***Activity Questions***	***Use***
1. Interview program leaders, staff, and beneficiaries to identify their judgments about what program successes should be sustained or expanded. →	- What successes do the interviewees identify? - Why are they considered successes? - Why do the interviewees think these successes warrant sustaining the program? - How are program planners planning for sustainability? Institutionalization? - What steps are being taken towards sustaining the program?	<u>Use Sustainability findings to:</u> - Determine whether staff and beneficiaries favor program continuation - Assess whether there is a continuing need/demand for sustaining the program's services - Determine how best to assign authority and responsibility for program continuation - Plan and budget continuation of the program
2. Review data on program effectiveness, program costs, and beneficiary needs to judge what program successes should and can be sustained. →	- What did the Effectiveness evaluation results indicate in terms of sustaining the program? - What were the program costs from its prototype and demonstration phases? - If there is a continued need for this program, do the findings indicate the program can and should be sustained? Why or why not?	
3. Interview beneficiaries to identify their understanding and assessment of the program's provisions for continuation. →	- Do beneficiaries understand whether the program will be continued or not? - Do they understand any potential changes to the program in its continuation?	
4. Periodically revisit the program to assess the extent to which its successes are being sustained. →	- Did the subsequent evaluation(s) demonstrate that successes have been sustained?	

Transportability

Purpose: Transportability evaluation assesses the extent to which a program has or could be successfully adapted and applied elsewhere.

Evaluation Activities	***Activity Questions***	***Use***
1. Engage the program staff in identifying actual or potential adopters of the program. →	- Who has adopted the program? - What was the adoption process? - What components of the program were adopted? - Who has expressed interest in adopting the program and why? - Who has visited/observed the program with an interest in adopting it?	Use Transportability findings to: - Assess the need for disseminating information on the program - Determine audiences for information on the program - Determine what information about the program should be disseminated -Gauge how well the program worked or could work elsewhere
2. Question potential adopters. →	Ask potential adopters to: - Review a description of the program and a summary of evaluation findings - Judge the program's relevance to their situation - Judge the program's quality, significance, and replicability - Report whether they are using or plan to adopt all or parts of the program	
3. Visit and assess adaptations of the program. →	- How have the adopters implemented the program? - How does this implementation compare to the original program? - Is the adaptation as successful as the original program? - How is the adaptation different from the original program? - What other observations can be made about the adaptation?	

Appendix B

Draft Evaluation Plan Template

[To be included in grant proposal submission; based on CCT Evaluation guidelines and informed by Kellogg Foundation evaluation guidelines][7]

Background and Context

1. Briefly summarize program purpose/goals, intended beneficiaries, and primary activities.

2. How will your planned activities affect progress in achieving program goals? Describe the main links between the activities and the goals that will lead to success in meeting your objectives.

   This question can be addressed by developing and describing a strategy map or logic model that illuminates what needs to happen to complete goals, including overcoming potential obstacles. It can also help in generating evaluation questions.

   Grant Lifecycle Stage:

   Refer to the Evaluation Guidelines document to determine your stage in the grant lifecycle. This may be revised based on discussions with the program officers. Check all that apply.

   ☐ Planning/Prototype ☐ Demonstration ☐ Maturity/Going to Scale

   Evaluation Guidelines Components:

   Based on the grant lifecycle stage identified above, check the appropriate evaluation components that you will address. Again, this may be revised based on discussions with the program officers. Check all that apply.

   ☐ Context ☐ Inputs ☐ Process ☐ Impact
   ☐ Effectiveness ☐ Sustainability ☐ Transportability

Purpose of Evaluation

What will the evaluation tell you about your program/initiative that you do not already know?

* Refer to the CCT Evaluation Guidelines to help identify potential uses for evaluation findings based on the appropriate lifecycle stage. For example, a pilot program might use evaluation findings to review appropriateness of program goals, refine program strategies, and monitor/improve implementation, among other evaluation goals.

[7] http://www.wkkf.org/Default.aspx?tabid=90&CID=281&ItemID=2810015& NID=2820015&LanguageID=0

Evaluation Questions

Using your strategy map/logic model and the CCT Evaluation Guidelines as a reference, identify appropriate evaluation questions with corresponding objectives and measures. This section should take the following format:

1. Goal #1
    a. Evaluation question “A”
        i. Measure 1
        ii. Measure 2
    b. Evaluation question “B”
        i. Measure 1
        ii. Measure 2

2. Goal #2
    a. Evaluation question “C”
        i. Measure 1
        ii. Measure 2
    b. Evaluation question “D”
        i. Measure 1
        ii. Measure 2

3. Goal #3 … (if needed)

Data Collection Methods

What key sources of information will be obtained to address each of the above evaluation questions and measures?

Disseminating Evaluation Results

Describe potential audiences (stakeholders) for the results of your evaluation and how the results will be shared. The Trust is one, but not the only audience to consider. Others within your organization, school leaders, teachers, community members, and other funders are all potential audiences for your report.

Evaluation Logistics

Briefly describe the following structures and procedures needed to complete your evaluation:

- Personnel coordinating the evaluation
- Program staff involved in evaluation activities
- Timeline for evaluation completion
- Evaluation budget
- Plans to use help from outside evaluation resources (consultants, collaborators, etc)

Appendix C

Logic Model Example Adapted from BSF Grantee Model

Project Goal: Provide instructional strategies for improving student reading and writing to every teacher in building.

Resource Inputs	Activities	Outputs	Outcomes		
			<u>Short</u>	<u>Medium</u>	<u>Long</u>
CCT/BSF grant	Identify literacy team members and set agenda	Monthly literacy team meetings held	Literacy team provides on-going feedback to teachers	Feedback and literacy tools consistently used by teachers	Student literacy skills improved
Staff literacy experts	Develop writing rubric and manual of style	Rubric and manual developed and shared with all staff	Rubric & manual provides feedback format	More teacher collaboration around literacy across content areas	Increase in student achievement results across content areas
Prior experience with PD provider	PD by teacher leaders	100% of teachers attend all professional development sessions	Teachers use rubric and manual to develop lessons	Model lessons developed and shared	Improved teacher efficacy

Assumptions	External Factors/Contingencies
Staff commitment to literacy strategy	Some teacher resistance due to preferences of prior practice or isolation norms, which may require one-on-one discussions

Appendix D
Draft Evaluation Report Template

ORGANIZATION

GRANT NUMBER

GRANT AMOUNT

PROJECT NAME AND DESCRIPTION (less than 100 words)

GRANT LIFECYCLE STAGE

Check all that apply from your evaluation plan.

☐ Planning/Prototype ☐ Demonstration ☐ Maturity/Going to Scale

EVALUATION GUIDELINES COMPONENTS

Check all that apply from your evaluation plan.

☐ Context ☐ Inputs ☐ Process ☐ Impact
☐ Effectiveness ☐ Sustainability ☐ Transportability

EVALUATION REPORT

Summary or Abstract:

Summarize key findings and implications of those findings for the program's future.

Background and Context:

This section is a repeat of the Background and Context section of the evaluation plan.

Main Goals:

Describe the main goals that were identified in the Background and Context section of the evaluation plan. Include any changes in the goals that have occurred since the original evaluation plan.

Main Questions:

List the main questions framing your evaluation and the measures used to answer the questions. Describe adjustments to any of the study questions or measures.

Results/Lessons Learned:

Report your evaluation findings for each evaluation question. Include references to the data you collected and how it was analyzed. Survey forms, interview questions and other sources of data collection should be attached to the report. Describe any adjustments to the program or the evaluation approach that were made during the course of your activities.

Conclusions/Next Steps:

Synthesize your findings and explain the implications of the findings for next steps in the program. For example, did the project proceed as expected? Based on your findings, do you need to adjust beneficiary groups or targets for your work? Do you need to change data collection procedures? What needs to be done to expand the program? Is another year of piloting in order with a different approach? Are there personnel issues?

This section should cover the big lessons for the program based on the evaluation results.

Appendix E

DESCRIPTIVE INFORMATION ON 2007–2008 BSF GRANTS

Project name	Grant amount	Group	Lifecycle	Project purpose
Implementing Student Literacy Initiatives Across the Curriculum	$18,654	Content/Literacy	Demonstration	Subsidize four full-day teacher institutes to implement the Literacy Across the Curriculum program. Teachers will work in cross-disciplinary teams and develop a writing rubric and manual of style. The program was designed with High Schools That Work.
Enhancing Teacher Competency with Technology Integration	$16,950	Content/Literacy	Planning/Prototype	Address skills teachers need to meet state teaching standards through professional development. Obtain instructional technology and train staff in use, specifically focusing on integrating technology into their classroom practice.
Teaching and Learning in a Literacy Community	$50,000	Content/Literacy	Planning/Prototype	Improve literacy skills. A 2-year literacy professional development partnership with Saint Xavier University to close achievement gap between school's pre-K–8 students and their more affluent peers. Literacy goals address instruction across the curriculum, collaborative coaching, and using assessment data to inform instruction.
Improving Student Writing by Implementing the Six Trait Writing Model	$45,775	Content/Literacy	Planning/Prototype; Demonstration	Implement the six trait writing model facilitated by the Great Source Education Group in three schools. Teachers would attend a 2-day workshop, receive follow up coaching, and an additional workshop.
Utilizing external research and best practices to review and improve one school's dual language program	$9,240	Data to inform instruction	Planning/Prototype	Implement a planning grant to learn about best practices and research on dual language instruction, analyze how that relates to the school's current dual language program, and produce a final report with recommendations. A task force of 6 faculty members and the principal will be facilitated by consultants from the Illinois Resource Center.
Academy Professional Development Program	$30,000	Data to inform instruction	Planning/Prototype; Demonstration	Continue a multi-year effort to improve the school's curriculum and enhance overall teacher performance. The two primary program elements are curriculum mapping and professional development and accountability.
Enlighten, Elevate, Achieve	$12,200	Data to inform instruction	Demonstration	Use *Curriculum Mapper* software and professional development, facilitated by consultants, to focus professional development and restructure curriculum around ACT's College Readiness Standards.
Professional Development Program A	$5,000	Data to inform instruction	Planning/Prototype; Demonstration	Support professional development agenda in (a) Curriculum and Learning Standards Development, (b) Support for English Language Learners (ELLs) and (c) Professional Development Seminars.

Project name	Grant amount	Group	Lifecycle	Project purpose
Professional Development Program B	$11,544	Data to inform instruction	Planning/Prototype; Demonstration	Utilize conferences and professional services to drive ongoing content oriented professional development in support of school's adoption of 13-point rubric to promote network-wide, ongoing, and programmatic professional development.
Professional Development in Technology	$20,000	Technology	Demonstration	Continue technology integration plan from 2006–2007, including installation of computer workstations in 15 classrooms, the completion of the laptop-based computer classroom, continuing education with workshops and seminars, and the launch of the new HTHS Technology Club and Computer Enrichment Program with one other school.
Technology Integration into Teaching and Learning	$85,000	Technology	Demonstration	Expand a 2006 grant to continue training teachers to integrate technology in an engaging and interactive manner.
Technology Coordinator	$30,000	Technology	Planning/Prototype; Demonstration	Support technology coordinator to work with faculty in order to develop their computer technology skills. Technology Coordinator will also consult and work with the principal and faculty at two other schools.
Promethean Technology Integration in the Daily Classroom	$85,000	Technology	Planning/Prototype; Demonstration	Training teachers at three schools to integrate Promethean technology into classroom instruction and to foster teacher collaboration through the technology.
An Integration of Technology – the Promethean Project 2008	$83,359	Technology	Planning/Prototype	Purchase Promethean ActivBoards and technology for 2 classrooms in each school; hold workshops on use of technology; foster learning community across schools; assess impact on student achievement to inform instruction and modify curriculum.
Educator coaching network	$88,500	N/A; general	Planning/Prototype; Demonstration	Provide professional development to schools in the BSF network, including mentoring, coaching, and selected professional development programs, and encourage peer collaboration.
Building Resources in Marketing/Fundraising	$75,000	N/A; general	Planning/Prototype; Demonstration	Hire a marketing consultant to work with thirty-six schools in the Patrons Program to create a marketing model; to create a catalogue of best practices to be shared across our system; and to act as a support person/mentor for these schools.

Appendix F

Example of Feedback on Evaluation Plan

Category	Prompts	Specific feedback
Background and context (and strategy map if available)	• Are beneficiaries and their needs clearly identified? • Are activities summarized? • Do activities proceed in logical manner to meet identified goals?	Patron school marketing/development directors and principals of those schools are the primary beneficiaries. How are these schools different from other elementary schools in the BSF network? The activities make sense. How many of the directors attend the bi-monthly Archdiocesan Development Council meetings? Are your activities aligned with the training in these sessions?
Grant lifecycle stage	• Is grant lifecycle stage realistic given goals and activities?	Yes, but I'm not sure about the sustainability component at this point. I'd be more concerned that the training and support provided is received by the directors and is influencing (supporting) their roles. Then when you know you have the right formula, the question can turn to how the efforts are expanded to other schools and sustained within the 20 Patron schools.
Purpose of evaluation	• Is utilization to the organization and its stakeholders clearly stated?	Stakeholders and potential changes to strategy are discussed.
Evaluation questions	• Are evaluation questions: ✓ specific ✓ measurable ✓ realistic ✓ connected to the goal • Will measures provide data to answer the evaluation questions? • Are measures realistic given resource constraints, capacity, time?	Have you thought of a macro-study question, e.g., How has the marketing/development director support been received by the directors and what influence has it had on their roles? Then the sub questions would logically follow because I think you are trying to determine the impact of your training and support, right? For your question on goals 1 and 2, you might want to frame it is a before/after question in order to capture change. For example: "How have student enrollment trends changed with the placement of marketing/development directors in Patron schools?" This would require getting trend data before and following it after. You might think about adding some open-ended forms of evaluation questions. They can reveal more information in a qualitative sense. So, for your question on accountability through metrics, you might ask: "How do marketing/development directors track/measure marketing and

		fundraising activities?" then "What reports do marketing/development directors produce on the results recruitment and fundraising efforts?" Using documents as primary data sources is a good start. How will they be measured for quality? Also, you might consider having informal interviews or feedback sessions from the directors to examine these questions as well as the perceived effectiveness of support they have received and suggestions for additional support. Since you include getting feedback from principals on how the directors have freed them for other leadership roles in schools, I suggest you frame an evaluation question to measure this, e.g.: "To what extent have principals been freed to work on other important leadership roles given the support provided by marketing/development directors?"

Appendix G

Evaluation Guidelines Project:
Grantee Interview Protocol, May 2008

We are interested in learning more about your program, your evaluation and communication with BSF and/or CCT about your activities, and your reactions to the evaluation guidelines material and professional development sessions this year.

1. In carrying out your evaluation activities as part of your grant with BSF, how was this year different from prior years?

    a. In the past, what communication did you have with BSF regarding evaluation?
    b. How did you typically structure evaluation activities in the past?
    c. What were the primary uses of your evaluation results in the past?
    d. How have your evaluation activities changed since participating in the evaluation guidelines pilot study with WCER and BSF?

2. How has your understanding the process and purpose of evaluation changed since participating in the pilot study?

3. I'd like to hear about the progress you have made with completing the evaluation of your program?
    a. How have you monitored and documented evaluation activities?
    b. What challenges did you face in your evaluation and how did you adjust to those challenges?
    c. How do you plan to use your evaluation results?

4. What are your reactions to the evaluation guidelines document?
    a. How did you use the guidelines during your evaluation planning and reporting?
    b. Was it useful? (probe for understanding, applicability, use to inform context, input, process, impact, effectiveness components)
    c. What changes would you suggest to make the document more user-friendly?

5. What are your reactions to the evaluation plan template?
    a. How did you use the evaluation plan?
    b. Was it useful? (probe for understanding, applicability)
    c. What changes would you suggest to make the document more user-friendly?

6. What are your reactions to the strategy map activity?
    a. How did you use the strategy map to inform your evaluation plan or activities?
    b. Was it useful? (probe for understanding, applicability, use to identify evaluation questions and key evaluation stages)
    c. What changes would you suggest to make the document more user-friendly?

7. What are your reactions to the evaluation report template?

   a. How did you use the evaluation report template? Were you able to address each area?
   b. Was it useful? (probe for understanding, applicability)
   c. What changes would you suggest to make the document more user-friendly?

8. Do you have any other suggestions on how the evaluation documents or information provided by BSF about evaluation requirements could be improved?