“Is our program working?” This is a key question in education today, particularly in this era of heightened accountability. A collaborative program evaluation model is an extremely useful way to answer this question when education organizations want to find out if their initiatives are achieving the intended outcomes, as well as why this is the case.

In the collaborative program evaluation model, the client (e.g., districts, states, public and independent schools, nonprofits, and foundations) works with the external evaluator to determine the questions that will be explored through the evaluation. They continue to work collaboratively to ensure that the context is understood, that multiple stakeholder perspectives are taken into account, and that data collection instruments are appropriate in content and tone. The model produces data that can proactively inform program implementation, provide formative information that supports program improvement, and offer summative information on the effectiveness of the program.

This PCG Education White Paper describes the benefits and essential elements of a collaborative program evaluation model. The paper is based on the experience of PCG’s research and evaluation team.
WHY HIRE AN EVALUATOR?

There are many reasons why an educational organization might choose to hire an evaluator to assess program effectiveness and determine if money is being well spent. An evaluation might be commissioned at one or more critical junctures of program implementation: when programs are first established, modified, or expanded; when stakeholders advocate for more information; when student outcomes do not meet expectations; or when a case for additional funding needs to be made. Skilled evaluators will apply a variety of data collection methods, approaches to analysis, and reporting techniques to respond to all of these situations.

Regardless of the type of educational program being evaluated, program directors and staff dedicated to continuous improvement want to know three things.

Evaluation Implementation

Is the program being implemented according to plan?

If the theory of action behind the project states that certain actions need to be taken and specific structures established or services delivered before results are achieved, it is important to make certain that these are in place and actually happening. Otherwise, there is little point in evaluating efficacy or impact.

For example, if teacher professional development were needed to show how to conduct Socratic circles in the classroom and the professional development did not take place, it would be a waste of resources to seek evidence of changed practice through classroom observations.

Evaluation of Efficacy

Is the program having the desired effect?

An important step in an evaluation is to document evidence that the program is having the intended medium-term effects that will presumably lead to the desired long-term outcomes. If the program is being implemented as planned, it is important to examine if the program has the power to produce the desired medium-term effects. For example, if regular meetings of trained data teams are supposed to result in changes in teacher practice, with the ultimate long-term result of increases in student achievement, it is important to check if changes in classroom practice are actually occurring. If adjustments to instruction that are aligned with evidence-based best practices are occurring, we would eventually expect to see improved student outcomes, assuming the program’s design is effective.

Evaluation of Impact

Is the program having the intended outcome?

Outcomes define what the program is designed to achieve. It is the role of the evaluator to examine whether or not the program has produced a change in particular outcomes over time. This analysis should occur in light of the program goals and the evaluation questions. For example: Does the well-run mentoring program result in fewer discipline referrals? Do students aspire to go to college in greater numbers after their school fully implements an arts-integrated program? Do students’ test scores increase as a result of using a computer-based math program with fidelity?

An external evaluator brings outside perspective and expertise to the task of assessing and reporting the degree to which educational programs are meeting the needs of students. In a collaborative evaluation, the client and the evaluator discuss the data and determine why this might be the case. Internal staff members may be too close to the work to be able to determine the impact of the program or they may not have time to step back and examine their work over time. Internal staff also may not have the expertise required to carry out a sound evaluation.

The evaluator can, based on the program, clarify the questions the client wants to answer, create an evaluation design that mixes and matches the appropriate data collection and analysis methodologies, design custom data collection instruments and approaches, and draw upon content expertise to provide valuable feedback and insight to those responsible for the programs.

THE BENEFITS OF COLLABORATIVE PROGRAM EVALUATION

Collaborative evaluation1 is a proactive evaluation model that enables program staff to engage in continuous program improvement. Specific benefits of the model include

- A customized evaluation design that reflects the nuances of the program being evaluated.
- An evaluation design that is flexible and adaptable to the purposes of the evaluation and to changes in program implementation over time.
- Increased validity of results.
- Greater buy-in among stakeholders with both the data-collection process and the evaluation findings.
- Development of program staff’s capacity to continue to monitor their progress toward program goals beyond the duration of the evaluation.
- Development of a culture of inquiry among program staff.
- Potential cost efficiencies.

Each of these benefits is described in detail below.

1 Note: There are a variety of other types of evaluation models that might be useful in certain contexts, such as randomized control studies, compliance audits, and descriptive studies. We contend that collaborative program evaluation is an excellent model for educational organizations to use when looking at program effectiveness.
THE POWER OF COLLABORATIVE PROGRAM EVALUATION

Address program nuances.
All evaluators should tailor evaluation services to the needs of each client (Patton, 2002). In the collaborative evaluation model, this is accomplished by evaluators working closely with program staff to identify evaluation questions and engage in an evaluation process that is attuned to the needs of program staff and stakeholders. As a result of the close knowledge built through collaborative program evaluations, such studies also guide program staff to identify and capitalize on external and internal program networks that they can tap to help them to achieve program goals (Fitzpatrick, 2012).

Flexible design.
In a collaborative evaluation, continuous communication at the outset between program staff and the evaluation team is essential for laying the groundwork for mutual understanding. Ongoing communication is also a key ingredient for ensuring that the evaluation plan continues to be relevant to the program. By communicating regularly about program developments and context, evaluators can make adjustments in the evaluation plan to accommodate changes in the program.

Increased validity of results.
Another benefit of working collaboratively with program staff in developing the evaluation is increased validity of the study. Because the evaluation team develops a high level of understanding of the program, data collection can be designed to accurately capture aspects of interest, and appropriate inferences and conclusions can be drawn from the data that are collected.

Greater buy-in for results.
Engaging an experienced outside evaluator alone increases the validity of the study and the credibility of the findings. The use of a collaborative program evaluation also improves buy-in for the study’s results from a variety of stakeholders. Staff members who actively participate in the evaluation better understand how the results can be used to facilitate program improvement, while administrators and other decision makers are more likely to have confidence in the results if they are aware that program staff helped inform elements of the evaluation study (Brandon, 1998).

Increased ability to monitor progress.
The evaluation team works with program staff to develop tools to measure desired outcomes of the program. Because tools are designed in collaboration with program staff, staff are better able to understand the purpose of the tools and what information can be gleaned from each. This makes it more likely that staff will feel comfortable with and use the instruments to collect data in the future to monitor ongoing progress, an added benefit to the client.

Development of a culture of inquiry.
Because use of evaluation results is a primary goal of collaborative evaluation, the evaluation team may also facilitate a process in which practitioners examine data on program implementation and effectiveness throughout early stages of the evaluation. This process of reviewing evaluation results can foster the development of a culture of inquiry among program staff and support the goal of continuous improvement.

Potential cost efficiencies.
There are several ways that a collaborative program evaluation can reduce costs in the short term and over time. There can be immediate cost savings because evaluation resources are tightly coupled with the program’s stage of development. The model can help avoid costly data collection strategies and analytic approaches when there is little to measure because the project is in a nascent stage of implementation.

Cost savings may also emerge over time because of program improvements based on formative feedback. Additional savings may be found as the evaluation team develops the internal capacity of program staff through their active participation in the design and execution of the evaluation. With increased capacity, the program staff can then continue the progress monitoring process by themselves.

The collaborative evaluation process occurs in three phases
1. Getting Underway: The phase where the theory of action is developed.
2. Full Engagement: The phase where designing data collection tools is undertaken, data collected, and findings reported.
3. Wrapping Up: The phase where an action plan is developed to use the evaluation results.

COLLABORATIVE PROGRAM EVALUATION IN PRACTICE
A collaborative program evaluation can employ a variety of approaches, but focuses on building a relationship between the evaluation team and program staff with the goal of building the capacity of program staff to use evaluation results and promote program improvement (O’Sullivan, 2012).

The process of a collaborative evaluation occurs in three general phases: (1) getting underway, (2) full engagement, and (3) wrapping up. While the phases appear linear, they are, in fact, dynamic and iterative as implemented throughout the evaluation process. Program staff and the evaluation team are engaged in a continuous cycle of dialogue to

• Build program and evaluation knowledge.
• Communicate about progress with the evaluation and developments in the program.
• Review evaluation findings and recommendations.
• Revisit, as necessary, evaluation questions and tools to ensure that they will generate key information needed for decision making and to make program improvements.
PHASE 1. GETTING UNDERWAY

By taking a collaborative evaluation approach to develop knowledge of the program and to design the evaluation, program development and implementation are also improved. During this initial phase of the evaluation, the evaluation team engages with program staff to answer a series of questions and clarify program details that will guide the evaluation design. In this phase, questions include:

- What are the official goals of the program?
- What steps are required to achieve these goals?
- What is the program’s current stage of implementation?
- What questions do we wish to answer about the program through the evaluation?
- What is the best way to measure the outcomes we’re interested in?
- What roles will the evaluation team and program staff play throughout the evaluation process?

The evaluation team seeks to understand program purposes, evolution, activities, functions, stakeholders, and the context in which the program operates. This is accomplished not only through a review of relevant program documents, but also through conversations with various stakeholders. A goal of this phase is to identify, with program staff, the theory of action that undergirds the program. That is, what do program staff believe needs to happen in order to get the results they seek?

CASE STUDY BRIEF

A theory of action assures the right questions are asked at the appropriate time.

In an evaluation study of a reading intervention program that was being implemented in several classrooms, a school district originally sought an evaluator to focus on student outcomes.

However, before an assessment of outcomes could occur, PCG staff worked with the district to develop an understanding of what full implementation of the intervention should look like by developing the theory of action.

The evaluator helped the program staff to see that if the program was not being implemented with fidelity (a mid-term outcome) and students had not had equal or quality access to the program, then there would be little sense in looking to see if students’ reading scores had improved (the long-term outcome).

Thus, the first evaluation questions focused on assessing fidelity of implementation: Are teachers using the program assessments? Are teachers working with the program for the recommended amount of time? Are teachers using the program technology and writing components of the program? Interim evaluation findings revealed that implementation was highly uneven across classrooms.

The evaluation team provided that feedback to the program staff who in turn made changes to improve the chance that the program would meet its long-term goals.

The theory of action is translated into a graphical representation called a program logic model. A logic model displays program inputs, as well as targeted medium effects and long-term outcomes. Review of the logic model can help the evaluators and program staff understand what assumptions are in place in the program, what the purpose of the program is, and what steps are needed to obtain the desired result (Dyson & Todd, 2010; Helitzer et al., 2010; Hurworth, 2008).

Stakeholders may not have formally articulated their theory of action and may be operating within several different theories depending on their vantage point (Weiss, 1998). The process of identifying an explicit model can focus the program staff as they develop and implement the program and allows the evaluation team to better match data collection and analyses to the project’s goals and objectives (Fear, 2007).

The evaluation team and the stakeholders work together to establish the logic model to develop a common understanding of the program. This process clarifies how the program components relate to one another and informs the development of questions to be answered in the study. Questions depend upon the program’s stage of implementation and might include:

- Is the program being implemented the way it was intended?
- What is the level of satisfaction of various stakeholders with the program?
- What is the effect of the program on teaching practices?
- What is the effect of the program on student achievement?

Agreement on and explicit identification of evaluation questions helps frame the scope of the evaluation, data collection activities, and the level of involvement in evaluation activities by the evaluator and other stakeholders. In this phase, evaluators may also conduct a review of existing research on similar programs to help support best practices and change processes within the program and to ascertain whether similar research has been conducted that can inform the study design.

Once questions are agreed upon and desired program outcomes are clear, program staff and evaluators collectively design the evaluation. This includes deciding on measures and types of data collection tools, data collection processes, timelines, plans for how to analyze formative data to drive program improvement, summative data that will be collected to demonstrate program impact, and when to report on findings.

The data collection processes (e.g., surveys, focus groups, observation checklists, interviews, student performance data) and analytical methods (e.g., qualitative, quantitative, mixed) proposed by the evaluation team will vary depending on the questions being asked; the scope of the evaluation; how much the program staff is able to assist with data collection given their time, skills, and interest; and the evaluation budget.

A collaborative program evaluation does not mean that program staff must participate in every evaluation activity. However, developing a collaborative evaluation design does require making explicit decisions about roles based on the feasibility, time, skills, and interest to participate in each phase (Corn et al., 2012). Because the evaluation design is developed in consultation with program staff, it is more likely to reflect an understanding of the nuances of the program and the concerns of stakeholders. Also,
THE POWER OF COLLABORATIVE PROGRAM EVALUATION

given that stakeholders participate in framing the evaluation design and process from the beginning, they are more likely to understand and use evaluation findings (Rodriguez-Campos, 2012).

**CASE STUDY BRIEFS**

**Developing surveys that represent core program ideas.**

For an evaluation of an arts program in a public school system, PCG evaluators, in conjunction with program staff, developed survey instruments by first establishing a list of categories of inquiry. Questions were developed within each category using program materials, a review of the literature, and interviews with staff, teachers, and artists at school sites. Survey items were reviewed with the program staff to confirm that the survey reflected their theory of action, tapped core ideas about program implementation and impact, and used local context-based language understood by all stakeholders.

**Optimizing data collection within the context of the program.**

In a PCG evaluation of a state-sponsored coaching program to support schools not meeting annual performance goals, the state was interested in whether coaches, placed in schools to provide content expertise, supported student growth. The evaluation team designed an evaluation plan to incorporate data collected from a number of sources including school visits, and worked closely with program staff to create the conditions for successful data collection. Based on experiences with a prior program, principals and teachers were anxious about the placement of coaches from “the state.” As a result, the coaches and state program staff had worked to build trust at their sites, which had to be maintained during the school visits conducted by the evaluation team. Therefore, it was collectively decided that a representative from the program would travel with the evaluation team and would participate in interviews with school staff as part of the evaluation. This clarified that the evaluation was supported by program staff, facilitated the collection of data, and put school staff at ease.

**PHASE 2. FULL ENGAGEMENT**

Collaborative evaluation enhances the quality of communication and level of trust with the client, which contributes significantly to the process of ongoing implementation, evaluation activities, and program improvement. After working in concert to articulate the theory of action and design the evaluation, the evaluation team and program staff are ready to fully engage with each other and the evaluation activities. Components of this phase are repeated in an ongoing cycle of data collection, analysis, reporting, and use of evaluation results. As the program and the evaluation evolve, this phase also includes periodically revisiting the evaluation plan to rethink evaluation questions in light of findings, any program developments that might influence the evaluation design or outcomes, and new questions that emerge.

In a collaborative evaluation, designing data collection tools is undertaken as a partnership between the evaluation team and program staff to ensure that the tools will appropriately measure the implementation and impact of a particular program (Lusky & Hayes, 2001). During this phase, evaluators and program staff come to consensus around the questions: What is the available evidence to answer the evaluation questions? How can we most effectively answer the evaluation questions? What is feasible to collect? Tools developed might include focus group or interview protocols, surveys, and observation checklists.

In addition to deciding what evidence can and should be collected, the evaluation team works collaboratively with program staff to optimize data collection opportunities. Program staff have knowledge of the climate and opportunities for data collection that will least interrupt the flow of daily program activities and will allow the evaluators to experience the program in an authentic way. Program staff can support data collection efforts by communicating needs directly to sites or staff. Involving program staff in data collection builds their understanding of both the evaluation process and the findings. The evaluation team shares evaluation and content expertise with program staff, and program staff share deep knowledge of their work and the context in which it is done. Collaborating in the data collection process builds staff capacity to conduct ongoing progress monitoring using these instruments beyond the end of the formal evaluation study.

Depending on the nature of the data collected, data analysis may follow a similarly collaborative process. Evaluators bring technical and conceptual expertise to the analysis of quantitative and qualitative data gathered, but it is through the expertise shared by program staff and collaborative dialogue that it becomes clear the types of analyses that will be most meaningful to program staff and other stakeholders. For example, individual schools and districts may wish to see their own survey or achievement data, whereas state administrators may be most interested in data aggregated by region or level of urbanicity. In addition, evaluators may bring findings to program staff as they emerge in order to collaboratively brainstorm possible explanations and additional analyses to pursue. For example, if a program designed to increase literacy achievement for all students seems to have a particularly large effect on students classified as English language learners, stakeholders may wish to delve more deeply into these data to more fully understand this finding.

Once the data have been analyzed and findings have been compiled, the evaluation team and the program staff must decide upon the most relevant ways (given multiple audiences) and intervals to report findings. Reports of findings should include an interpretation of those findings and recommendations appropriate to the specific context of the program (Poth & Shulha, 2008). Ideally, the reporting schedule should be arranged so that the findings can both inform the ongoing actions of program staff and enhance decision making by stakeholders. The evaluation team may agree to provide a set of interim reports or presentations that help program staff reflect on implementation and impact, including a formal written report with associated materials (e.g., an executive summary, a presentation, or a set of documents tailored for specific stakeholders).
A collaborative program evaluation process ensures that evaluation reports and related materials are tools that program staff can use to share information about their program with internal and external stakeholders. These tools can be used by staff to build a wider understanding of the program’s implementation, efficacy, and impact, to share context-sensitive recommendations for greater stakeholder involvement and program improvement, and to provide key information about the program to potential funders.

The evaluation team and the stakeholders collaboratively review the evaluation findings. This review reinforces understanding of the results and methods used to obtain them. Based on the review, the evaluation team and the program staff generate additional questions raised by the data and clarify next steps for the evaluation. For example, program staff and evaluators might consider: Given changes in program personnel and more rapid than planned expansion, what measures are most salient to assess progress toward goals? Is there a need for additional or alternate data collection tools?

**CASE STUDY BRIEF**

**Formative feedback supports attainment of project goals.**

In the evaluation of the same arts-integration program mentioned earlier, the annual evaluation report included information about program selections made at each school over time.

In considering which aspects of the evaluation report would be most useful to share with school-based site coordinators, program staff decided to share school-by-school and district-level results.

Sharing school and district data provided program staff with a platform from which to revisit a discussion of overarching program goals that call for distribution of program selections across multiple art forms and types of arts experiences.

Consulting annual program selections and district-level results helped school-based staff ensure that future program selections represent these goals.

**PHASE 3. WRAPPING UP**

In a collaborative evaluation model, the final phase lays the groundwork for program staff to build upon and continue to use evaluation results, even after the conclusion of the evaluation contract. Questions related to program implementation that may be answered during this phase include: How can we make best use of the evaluation findings? Based on current implementation, what steps need to be taken to increase fidelity? How do we create conditions to expand and/or continue our successes? Additional questions specific to the evaluation may also be considered near the conclusion of the evaluation period such as: What are our evaluation needs going forward? What infrastructure and leadership will support ongoing data collection and use of evaluation results?

At the core of the collaborative program evaluation model is the use of evaluation results, not only to understand program impact and inform decision making, but also to improve program implementation and student outcomes (Cousins & Whitmore, 1998; O’Sullivan, 2012). Use of evaluation data traditionally has not been part of the evaluation cycle. In many types of evaluations, the assimilation of evaluation results and development of a plan based upon them has often been left to the evaluation client. In a collaborative program evaluation, the evaluators may facilitate data-informed action planning to help program staff develop a plan to implement recommendations from the evaluation. Often, this is explicitly built into the process from the beginning.

By working together throughout the evaluation process, both partners develop a deeper understanding of how the program operates and what impact is anticipated. Consequently, the evaluation team is better equipped to provide actionable recommendations relative to areas that need improvement.

The evaluation team can also plan with the project staff how they might continue to use the data collection tools or evaluation processes developed for their program to continue to track and monitor the ongoing implementation and the effectiveness of program improvements. Supporting the program staff to develop this capacity facilitates implementation of recommendations and subsequent evaluation of how the recommendations are implemented.

Involving program staff with the evaluation design, data collection, review of evaluation results, and discussion of recommendations can position staff to continue the cycle of inquiry and action initiated by the evaluation. In fact, several studies have demonstrated how a collaborative program evaluation process can help a school develop and sustain a learning culture (Hoole & Patterson, 2008; Suárez-Herrera, Springett, & Kagan, 2009). Iterative review of evaluation results and recommendations builds what Fitzpatrick (2012) calls “evaluative ways of thinking—questioning, considering evidence, deliberating” into the life of an educational organization.

**CASE STUDY BRIEF**

**Facilitated action planning based on evaluation results builds a learning culture.**

In the evaluations of several Smaller Learning Community (SLC) grant recipients, the evaluations were structured to provide program staff with frequent data to inform progress toward program goals.

This included survey and interview data related to the extent to which the program created a learning environment conducive to student growth, increased the achievement of all students, and established a schoolwide culture that supported more personalized learning. Achievement data were also examined.

Coaching from the evaluation team was built into the evaluation to develop staff capacity to make data-informed instructional decisions.

Coaching supported the staff to write meaningful objectives, delineate action-oriented steps, and identify success indicators as part of their plan to advance the program.
THE POWER OF COLLABORATIVE PROGRAM EVALUATION

CONCLUSIONS

The primary reason that program staff can trust the findings of a quality collaborative program evaluation is because they know that the evaluator understands their context and their concerns and will work with them to achieve their goal of continuous program improvement.

As described in this white paper, the benefits of collaborative program evaluation include:
- an evaluation design based on sound principles that reflects the nuances of the program being evaluated;
- an evaluation design that is flexible and adaptable to the purposes of the evaluation;
- increased validity of results;
- greater buy-in among stakeholders in both the process and results;
- development of a culture of inquiry among program staff;
- development of program staff's capacity to continue to monitor progress toward program goals beyond the duration of the evaluation period; and
- potential cost efficiencies.

The collaborative program evaluation model allows the evaluation team and program staff to stand shoulder-to-shoulder in determining how to improve program implementation and effectiveness, thereby increasing the probability of improved student outcomes. In this type of evaluation, evaluators apply appropriate data collection and methods of analysis to determine whether the program is having the desired impact and provides recommendations for program improvements. While a collaborative program evaluation requires an ongoing commitment by all parties, it also produces high value to stakeholders and greatly increases the likelihood that educational programs will meet their intended goals and objectives.

REFERENCES


ABOUT THE AUTHORS

Christine Donis-Keller is a member of the research and evaluation team at PCG Education. Christine has worked in the fields of philanthropy and education and evaluation research for more than 20 years. Trained in the sociology of education, Christine has led and participated in evaluation projects, both large and small, within the education sector across the U.S. Her current work includes a collaborative program evaluation of Hartford Performs, a districtwide arts initiative, and an evaluation of the Tennessee Academic Specialists Program, a statewide program designed to support schools in need of improvement. She recently completed a national research project examining data use by teachers, schools, districts, and states for a major education foundation. Christine has authored research reports on theme high schools and is a noted expert on the four-day school week. She specializes in the policy implications of evaluation findings. Christine can be reached at cdonis-keller@pcgus.com.
Julie Meltzer, Ph.D., is Strategic Education Advisor for PCG Education. In the past 12 years, Julie has designed, led, or contributed to many educational evaluations and program reviews. Her belief in the collaborative program evaluation model comes from the power she has seen when it is used with districts and nonprofit organizations whose interest is in continuous program improvement as well as rigorous documentation and analysis of outcomes. Julie is a sought-after speaker about adolescent literacy, 21st century skills, transformational leadership, and systems thinking. She is the author and co-author of many books, articles, and professional development materials. Julie can be reached at jmeltzer@pcgus.com.

Elizabeth Chmielewski is a consultant at PCG Education. She focuses on managing PCG’s evaluation projects and providing research assistance. Currently, Elizabeth manages all project logistics for the Hartford Performs program evaluation, a three-year mixed methods study that examines the implementation and effects of a districtwide integrated arts program, as well as for the Tennessee Academic Specialists Program evaluation, a statewide effort to support high priority schools. She also supports the development of data collection tools, facilitates focus groups and interviews, and contributes to report writing. Previously, Elizabeth co-authored Investing in Community Colleges of the Commonwealth: A Review of Funding Streams and provided data analysis for several additional research projects. Elizabeth can be reached at echmielewski@pcgus.com.

We would like to thank Dr. Michelle LaPointe for her substantive contributions to the original paper on which this White Paper is based. Thanks to our internal reviewers and valued colleagues Mary Ellen Hannon, Keith Holt, Dennis Jackson, Nora Kelley, Erin MacIntire, Michelle Simmons, Stephen Smith, and Nicola Williams.

We always like to hear from readers. If you would like to provide feedback on this White Paper, please email your feedback to jmeltzer@pcgus.com. Please include your name, title, organization/district and state. Thank you.