

Brief Research-to-Results

Child TRENDS[®]

...information for program providers and funders who seek to understand what performance management is, how it differs from evaluation, and why both are critical.

Publication #2011-02

January 2011

PERFORMANCE MANAGEMENT AND EVALUATION: WHAT'S THE DIFFERENCE?

Karen E. Walker, Ph.D., and Kristin Anderson Moore, Ph.D.

OVERVIEW

In previous research briefs, Child Trends has provided overviews of various forms of program evaluations: experimental, quasi-experimental, and process.¹ This brief provides information on performance management—the ongoing process of collecting and analyzing information to monitor program performance—and its relationship to other forms of evaluation.

BACKGROUND

Funders increasingly require managers of social programs to provide information about their programs' operations, including services provided, participation levels, outcomes, and effectiveness. As a result, organizations need the capacity to collect significant amounts of information. How managers use information once it has been collected varies across organizations. For some managers, the usefulness of information ends once they have met their funder's requirements. Such managers may be collecting performance *measures*, but they are not engaging in performance *management* as we describe it below.

Other managers regard data collection and analysis as ongoing parts of their jobs that enable them to assess aspects of their programs and modify activities as needed. Recently, the federal government has emphasized tying funding levels to evidence-based programs, a strategy that is likely to benefit programs in which managers use varied types of data effectively to monitor and improve their programs. Thus, it has become more and more essential to use information for program improvement—an undertaking that is often called *performance management*. In addition, outcomes evaluation is increasingly expected. However, many nonprofit programs, unaccustomed to extensive data collection and analysis, are unsure what data to collect, how to use their data, and how performance management differs from evaluation.

WHAT IS PERFORMANCE MANAGEMENT?

The U.S. Office of Personnel Management defines performance management as “the systematic process by which an agency involves its employees, as individuals and members of a group, in improving organizational effectiveness in the accomplishment of agency mission and goals.”² The process is used for program or organizational improvement and supported by program managers who have the authority to modify program components. To carry it out, managers and other staff members collect information about the frequency, type, and (ideally) quality of services, along with information on participant satisfaction (or dissatisfaction) with and use of services and participants' outcomes. Managers then use the information to develop a better

understanding of their programs’ strengths and challenges, which helps in identifying solutions to improve program operations.

WHAT IS EVALUATION?

In this brief, the term “evaluation” refers to one of several types of assessments of social programs. For example, *implementation evaluations* are used to assess a program’s operations, and various types of *outcomes evaluations* are used to assess the effect of a program on participants and the program’s cost-effectiveness.

HOW IS PERFORMANCE MANAGEMENT SIMILAR TO EVALUATION?

Performance management and program evaluation share key features. Both rely on systematically collected and analyzed information to examine how well programs deliver services, attract the desired participants, and improve outcomes. Both use a mix of quantitative and qualitative information. *Quantitative information* is data displayed in numbers that are analyzed to generate figures, such as averages, medians, and increases and decreases. *Qualitative data* include information obtained through interviews, observations, and documents that describe how events unfold and the meanings that people give to them. In both performance management and program evaluation, qualitative and quantitative data are used to examine relationships among people, program services, and outcomes.

In general, performance management is most similar to *formative evaluation*, a particular type of evaluation that is intended specifically to provide information to programs about operations (such as whether the program meets high standards and whether targeted groups enroll). Although this brief distinguishes between performance management and evaluation in general, it focuses predominately on the distinctions between performance management and formative evaluation, sometimes called implementation evaluation.

<p>Information collected in Performance Management and Evaluation is similar and includes:</p> <ul style="list-style-type: none"> ▪ Participant characteristics ▪ Services and activities provided ▪ Attendance and retention

HOW IS PERFORMANCE MANAGEMENT DIFFERENT FROM EVALUATION?

Despite their similarities, performance management and evaluation differ from each other with respect to the purposes of collecting information, the timing of data collection, the people primarily responsible for the investigation, and how benchmarks are derived and used.

Performance Management	Evaluation
<i>Why is program information collected?</i>	
To ensure that the program is operating as intended	To understand current program performance and provide recommendations for program improvement
To plan and guide improvements if the program is not operating as intended or not producing desired outcomes	To describe program operations

Performance Management	Evaluation
	To assess program effectiveness or impact
	To assess the costs and benefits of the program
	To explain findings
	To understand the processes through which a program operates
	To assess fidelity of implementation
<i>Who is the intended audience?</i>	
Program managers and staff	Funders, policy makers, external practitioners
<i>When is information collected?</i>	
Throughout the program's life	Once or periodically
<i>Who leads the investigation?</i>	
Program staff	External evaluator
<i>How is progress measured?</i>	
Benchmarks are established for key measures and program progress is measured against them.	Progress is measured by increases and decreases in desired outcomes.

WHY IS PROGRAM INFORMATION COLLECTED?

In *performance management*, programs collect information to make sure that the program is operating as intended. If it is not, then the information helps managers decide among possible explanations for the program's challenges in meeting expectations. For example, information collected at enrollment about participant characteristics, such as school grades, indicates if an after-school program that targets academically at-risk youth is recruiting its intended population, but it does not explain why the program is or is not doing so. For that, the manager must collect additional information about the reasons that the targeted population is—or is not—participating. Taking a close look at how the achievement of youth who expressed interest in the program initially compares to the achievement of those who actually enrolled might provide some important clues. Perhaps the program's initial recruitment efforts resulted in attracting fewer high-risk youth than desired, or perhaps the program attracts both at-risk and low-risk youth, but it has trouble maintaining enrollment among the at-risk population. Managers need information about program participants, enrollments, and program use to monitor progress and make informed decisions.

In performance management, managers may collect information on participant outcomes, but such outcome data cannot be used to make a strong case for a program's effectiveness. Instead,

this information should be used in conjunction with information about participants and services to assess program operations. For example, if one group of participants is achieving desired outcomes, a program manager might ask such questions as, “How do these groups differ? Is one group already doing better when it begins the program? Do different staff members work with this group?” Such questions can help to identify ways to strengthen recruitment efforts or staff training.

To make a strong case for a program’s effectiveness, evaluation is necessary.

In *evaluation*, programs collect information for multiple reasons, depending on the type of evaluation. Outcomes evaluations summarize participant outcomes without trying to establish whether the program caused observed changes, whereas impact evaluations try to establish cause-and-effect through high-quality randomized controlled studies.³ Various types of implementation evaluations provide information that describes features of effective programs, and explain why programs may not reach their intended goals. Formative evaluation, one type of implementation evaluation, may be used to identify operational challenges that need improvement, just as performance management does. However, formative evaluation often has a broader goal, which is to identify common challenges and opportunities that programs face to prepare funders, policy makers, and practitioners for future work.

WHO IS THE INTENDED AUDIENCE?

The audiences for performance management and evaluation differ. The results of *performance management* are intended for the people within an organization who manage and run the program. In contrast, the results of *evaluation* are usually intended for external audiences, including funders, policy makers, and practitioners.

This difference in intended audience means that information collected solely for program improvement is not typically subject to approval by an Institutional Review Board (IRB). These boards evaluate proposed research to ensure that it meets ethical standards. These standards include respect for persons (by upholding the requirement that participation in research is voluntary, for example); beneficence (i.e., protecting research participants from harm and ensuring that the benefits of the research outweigh its risks); and justice (spreading the burden of research across different types of groups instead of selecting those who are vulnerable).⁴ IRB review, however, becomes important when an individual’s information is collected with the intent of incorporating it into more generalized knowledge that is disseminated to the broader community.

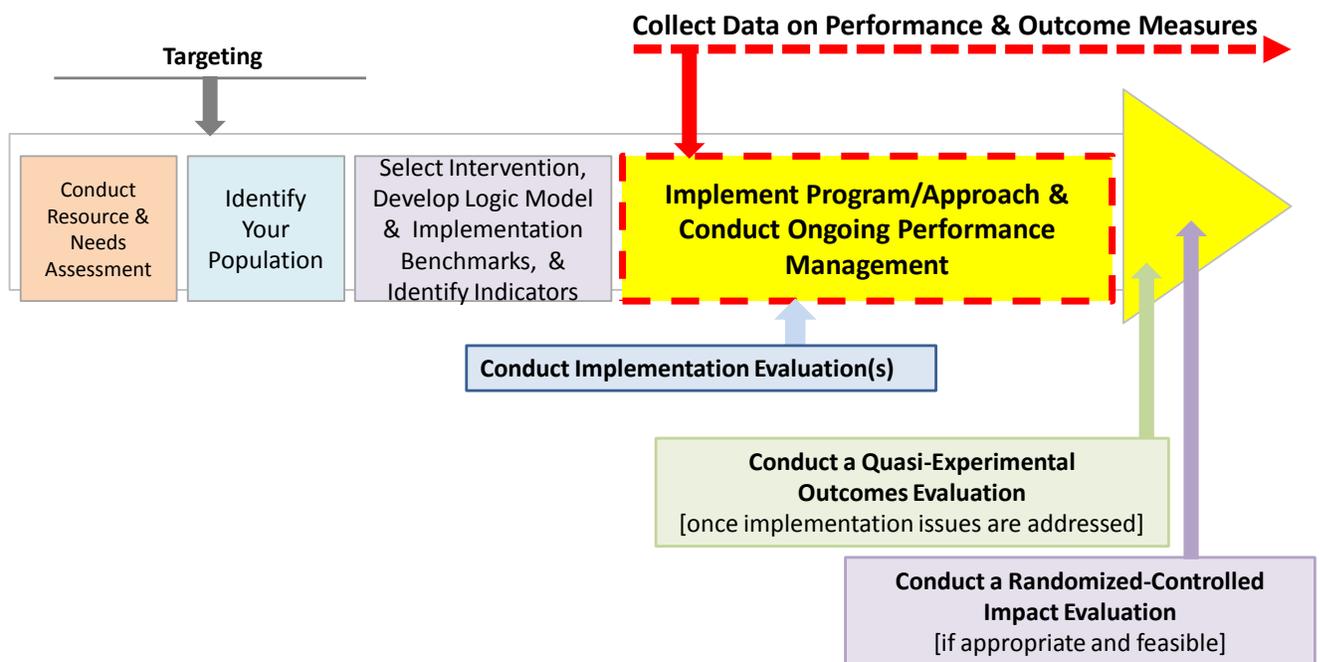
WHEN IS PROGRAM INFORMATION COLLECTED?

As indicated in Figure 1, information for *performance management* is collected throughout a program’s life on an ongoing basis. When a program is first developed and implemented, information such as participant characteristics, participation rates, and participant outcomes provide useful insights into critical questions. Among these critical questions are whether the program is serving its targeted population; whether participation rates are at the desired levels; and whether the participants’ outcomes are moving in the desired direction. Knowing the answers to these questions allows program staff to ask follow-up questions, such as “Why aren’t our participation rates as high as we would like?” and “What can we do to increase them?”

Even mature and stable programs need to monitor their progress. Changes in participant characteristics may result in falling attendance rates in a program, suggesting the need to make corresponding changes in the program. For example, if a program is seeing an increase in the ratio of immigrant to non-immigrant participants, hiring staff who can work effectively with the immigrant population may serve to boost program attendance and retention. Also, changes to public funding may lead managers to make small changes to their program that, over time, may add up to big and unintended changes in the experiences that the program participants have in the program.

Unlike performance management, *evaluation* is conducted periodically, and its timing depends on program maturity, the type of evaluation desired, and the organization’s need to have an outsider’s perspective. When programs are young or undergoing significant changes, formative evaluation may be appropriate after program staff members and managers have gone as far as they can in assessing and improving their own performance. Once evaluators are assured that the program is ready for an investigation of its outcomes, then impact evaluations may be conducted—along with implementation evaluation to explain findings.

Figure 1: Becoming Performance Driven



WHO LEADS THE INVESTIGATION?

Program staff or outside consultants generally conduct the research intended for *performance management*. To gain a better understanding of how to use information to improve program operations, the organization running the program should collect and analyze the data. This approach permits the organization's managers to ask questions and collect information about operational challenges in a timely manner.

Evaluations led by independent researchers whose fees are paid by funders (such as a government agency or a foundation) may not be as useful to program staff as efforts that are internally funded and led because the program has relatively little authority to determine the research questions or the methods used to collect and analyze data. Also, program staff may be anxious about the role of external evaluators, fearful that funders will make decisions to continue or discontinue support for a program based on evaluation findings. This anxiety may lead to distrust about the evaluation's findings, particularly if some of the findings are negative. Program staff members may focus on defending their program instead of asking what they could do to improve it. And, finally, program staff may distrust evaluators' conclusions because the staff does not fully understand how these conclusions were reached.

In contrast to program managers and staff, funders and policy makers usually think that independent evaluators provide more credible information than program staff or evaluators hired by programs do because there is no obvious conflict of interest. Thus, many evaluations are conducted by independent evaluators.

HOW IS PROGRESS MEASURED?

A fundamental difference between performance management and evaluation involves setting benchmarks to assess progress on indicators. A benchmark establishes a standard against which implementation progress is measured. Typically, benchmarks are expressed as numbers or proportions. For example, a summer literacy program may determine that its participants require six weeks of involvement to prevent summer learning loss. Six weeks, therefore, becomes the benchmark against which the program measures its progress in delivering services.

To a very large degree, the evidence that programs use for establishing benchmarks may be sparse or weak because it is based on limited research and program experience. For example, if six-week summer reading programs have been found to be effective in preventing summer learning loss in some situations, then a program manager may use six weeks as a benchmark, even if further research might determine that five weeks would be sufficient. Alternatively, program staff may set a benchmark based on assumptions about what participants need to benefit from the program. Such was the case with Big Brothers/Big Sisters mentoring program affiliates, which set initial program standards prior to a rigorous evaluation in order to communicate expectations to mentors and ensure that they had sufficient time to devote to the relationship.⁵ After experimental evaluations were completed, the benchmarks and standards were updated. Setting benchmarks for performance serves a critical operational purpose: It allows program managers to determine if activities are meeting standards and make decisions accordingly.⁶

Whereas program staff and managers usually are willing to use both experience and previous research to set benchmarks to assess their programs' progress, evaluators hesitate to set a benchmark without sufficient and strong data. An evaluator's work is judged according to

standards for data collection and analysis set by other evaluators and researchers. From an evaluator's perspective, carefully drawn conclusions are critical to advancing knowledge, which is developed incrementally and over time. Often, the most an evaluator can say is that the program is or is not effective based on a statistically significant minimum detectable effect. There are rarely opportunities to compare one benchmark with another. However, once an evaluation is complete, its findings may provide program operators with information that helps them refine their benchmarks for more effective performance management.

The practical result of these different approaches to setting benchmarks can be tension between program managers, who want evaluators to tell them what the standards should be, and evaluators, who say they do not have that information. If available at all, that type of information tends to be scattered unevenly across the evaluation literature.

CONCLUSION

Both performance management and evaluation can provide very useful information, but they are not interchangeable. Each has its own goals, standards, and criteria for meeting them.

Performance management aims to ensure that social programs operate as intended. It requires ongoing, internal data collection and analysis, flexibility to ask a variety of questions, and the capacity to use experience and the literature to set program standards and benchmarks.

Evaluation is intended to provide information to a broad set of stakeholders—funders, other practitioners, and policy makers—to advance knowledge in the field. It requires a clear set of research questions, the most rigorous design possible given programmatic constraints, and careful, time-consuming data analysis.

Both are critical to operating successful social programs, but it is important to note that, while performance management is a necessary task for every operating social program, evaluation is not. If a program model has evidence of effectiveness through a rigorous previous evaluation and program standards have been defined (either through the evaluation or by the program developers), then further evaluation may not be necessary. Performance management, however, remains a critical ongoing task to ensure that the program is meeting its standards. And if a program has never been evaluated, ensuring that the program is operating as intended is a good way to prepare for any future evaluation.

ACKNOWLEDGEMENTS

We would like to thank Ingvild Bjornvold for insightful comments that strengthened this brief.

Editor: Harriet J. Scarupa

REFERENCES

- ¹ Bowie, L., & Bronte-Tinkew, J. (2008, January). Process evaluations: A guide for out-of-school time practitioners. (*Research-to-Results* Brief) Washington, DC: Child Trends. Moore, K.A., & Metz, A. (2008, January). Random assignment evaluation studies: A guide for out-of-school time practitioners. (*Research-to-Results* Brief). Washington, DC: Child Trends. Moore, K.A. (2008, January). Quasi-experimental evaluations. (*Research-to-Results* Brief). Washington, DC: Child Trends.
- ² "Performance Management Overview," downloaded 7/19/10 from http://www.opm.gov/hcaaf_resource_center/assets/Ropc_tool3.pdf.
- ³ Allen T., & Bronte-Tinkew, J. (2008, January). Outcome evaluations: A guide for out-of-school time practitioners. (*Research-to-Results* Brief) Washington, DC: Child Trends. Bowie, & Bronte-Tinkew (2008). Moore, K.A., & Metz, A. (2008, January). Random assignment evaluation studies: A guide for out-of-school time practitioners. (*Research-to-Results* Brief). Washington, DC: Child Trends. Moore, K.A. (2008, January). Quasi-experimental evaluations. (*Research-to-Results* Brief). Washington, DC: Child Trends..
- ⁴ Bronte-Tinkew, J., Allen, T., & Joyner, K. (2008, February). Institutional review boards (IRBs): What are they, and why are they important? (*Research-to-Results* Brief). Washington DC: Child Trends.
- ⁵ Tierney, J.P., & Grossman, J.B., with Reisch, N.L. (1995). Making a difference: An impact study of Big Brothers/Big Sisters. Philadelphia: Public/Private Ventures.
- ⁶ Hunter, D. E.K., & Koopmans, M. (2006) Calculating program capacity using the concept of active service slot. *Evaluation and Program Planning*, 29, 186-192.

SPONSORED BY: The Atlantic Philanthropies

© 2011 Child Trends. *May be reprinted with citation.*

4301 Connecticut Ave, NW, Suite 350, Washington, DC 20008, www.childtrends.org

Child Trends is a nonprofit, nonpartisan research center that studies children at all stages of development. Our mission is to improve outcomes for children by providing research, data, and analysis to the people and institutions whose decisions and actions affect children. For additional information on Child Trends, including publications available to download, visit our Web site at **www.childtrends.org**. For the latest information on more than 100 key indicators of child and youth well-being, visit the Child Trends DataBank at **www.childtrends.org/databank**. For summaries of more than 500 experimental evaluations of social interventions for children, visit **www.childtrends.org/LINKS**.