Integrating Evaluative Capacity into Organizational Practice

Mission
Strategy
Finance
Fund Raising
Evaluation
Program Development
Marketing and Communications
Technology
Staff Development
Human Resources
Collaboration

ASK ?S
GATHER DATA
ANALYZE DATA
USE FINDINGS

A Guide for Nonprofit and Philanthropic Organizations and Their Stakeholders

Anita Baker, Beth Bruner
The Bruner Foundation, 2012

BRUNER FOUNDATION

EFFECTIVENESS INITIATIVES

Integrating Evaluative Capacity into Organizational Practice

BRUNER
FOUNDATION

EFFECTIVENESS
INITIATIVES

A Guide for Nonprofit and Philanthropic Organizations
and Their Stakeholders

Anita Baker, Beth Bruner
The Bruner Foundation, 2012

Acknowledgments

This guidebook was developed by Anita Baker and Beth Bruner based on 16 years of collaboration targeted at helping organizations build evaluative capacity—the capacity to do evaluation, to commission evaluation, and to think evaluatively.

The authors' efforts to increase knowledge and use of participatory program evaluation among nonprofit service providers and grantmakers began in 1996 with a pilot initiative, the Rochester Effectiveness Partnership (REP). Evaluations of each phase of the initiative provided clear evidence that through REP, participants learned about, could do, and could commission better evaluations of their programs. As the initiative evolved, alumni study groups were created to help participants sustain their evaluation capacity over time. Additions were made to the curriculum to address how evaluation skills could extend or "ripple" beyond participants into other programs, departments, or organizations. At the conclusion of the REP in 2003, Baker, Bruner, and a number of the former REP partners were curious about how increased evaluation skills might be useful beyond the development, implementation, and improvement of programs. In other words, could evaluation play a role in human resources or technology acquisition or marketing? A group of 12 organizations spent one year (2003–2004) thinking about transferring evaluation skills into multiple organizational areas through the use of what they defined as Evaluative Thinking. The Evaluative Thinking in Organizations Study (ETHOS) documents this effort.

The Hartford Foundation for Public Giving's Building Evaluation Capacity (BEC) began in 2006 and continues to this day. BEC builds upon the REP work, is also delivered by Anita Baker, and contributes to the work of integrating increased evaluation skills and organizational evaluative practice.

It is the belief of the authors and our many partners over time that building a culture of evaluation—one in which evaluation skills are firmly embedded and used for development, improvement, and decision making about programs, strategies, and initiatives—is fundamentally important in creating a learning organization, a data-driven organization, an organization sensitive to making informed decisions in a timely way. We also believe that evaluation skills are highly transferrable beyond evaluating an organization's programs, strategies, and initiatives. **Evaluative Thinking is just that—a way of thinking, a process of acquiring and using relevant data in any aspect of an organization's work.**

That is what this guidebook is about—**EVALUATIVE CAPACITY**—the integrating of evaluation skills and evaluative thinking into everyday organizational practice to ensure not only stronger programs but also stronger, more effective organizations better able to deliver on their missions. The contents of this guidebook were influenced by the former REP partners, the former and current BEC partners, as well as the work of Michael Quinn Patton, Paul Connolly, Peter York, Paul Light, Hallie Preskill, the Grantmakers for Effective Organizations (GEO), Mark Kramer, and others who regularly share their insights through American Evaluation Association and GEO blogs and publications.

A special thanks to our reviewers whose suggestions helped to frame this guidebook: independent evaluators Victoria Dougherty and Jodi Paroff, grantmakers Amy Studwell from the Hartford Foundation for Public Giving and Rebecca Donham from the MetroWest Health Foundation, and our nonprofit sector representatives, Executive Directors Tracy Hobson from the Center for Anti-Violence Education (CAE) in Brooklyn, New York, Ann Marie Cook from Lifespan of Greater Rochester, Inc., and Allison Overseth, Partnership for After School Education (PASE), New York City.

Foreword

Preliminary work on the Rochester Effectiveness Partnership (REP) began in 1995. Between 1996 and 2003, over 150 people in Rochester, NY participated directly in the initiative whose outcome was to build individual and organizational evaluation capacity using a participatory model.

A second interactive initiative, Evaluative Thinking in Organizations Study (ETHOS), was conducted in Rochester throughout 2004 with the purpose of understanding more about the relationship between increased evaluation capacity and the use of evaluative thinking—thinking grounded in evaluation skills but focused beyond the program level.

Three publications were developed from this work. All are available for download on the **Bruner Foundation**, website: http://www.brunerfoundation.org/. Those who want a hard copy can make a request through the website as well.

> *Participatory Evaluation Essentials: A Guide for Nonprofit Organizations and their Evaluation Partners*. Published in 2004 and updated in 2010, this guidebook provides a detailed curriculum for building participatory evaluation capacity. An automated version of the guidebook is available in PowerPoint format.
>
> *Participatory Evaluation Essentials: A Guide for Funders and their Evaluation Partners*. Published in 2004, this guidebook provides a detailed curriculum for use by funders. The funder materials were updated and enhanced as a series of PowerPoint guides.
>
> *Evaluation Capacity and Evaluative Thinking in Organizations*. Published in 2006, this monograph includes the history behind the development of the Evaluative Thinking Assessment tool as well as examples of its practical use in an organization. An example report of Evaluative Thinking Assessment results is also available.

This publication, *Integrating Evaluative Capacity into Organizational Practice,* was developed in response to the continuing need expressed by nonprofit trainees to further assess and operationalize evaluative thinking. It extends information first provided in 2006 in a series of short, electronic articles called *Evaluative Thinking Bulletins*. The guidebook is intended to answer the following questions:

- What is evaluative thinking, why is it important, and how can it be assessed?
- What is the relationship between evaluation and evaluative thinking?
- How can organizations use evaluative thinking in multiple areas of their work?
- What evaluation skills are critical to evaluative thinking and how can organizations enhance evaluation capacity?
- How can organization board members enhance and learn from evaluation?
- How is evaluative information packaged and distributed?
- How can an organization sustain its evaluative capacity?

These materials are for the benefit of any 501(c)(3) organization and may be used in whole or in part provided that credit be given to the Bruner Foundation. They MAY NOT be sold or redistributed in whole or in part for a profit.

Contents

1. Evaluation* and Evaluative Thinking

Key Definitions

Evaluation is a process that applies systematic inquiry to program management, improvement, and decision making. Evaluation is also used to assess the status or progress of a strategy (i.e., a group of meaningfully connected programs, not just the simple aggregation of multiple programs) or an initiative (a grouping of strategies).

Evaluation Capacity is the ability of staff and their organizations to do evaluation. Because evaluation is systematic, it involves proficiency in a particular set of skills.

- Asking questions of substance
- Determining what data are required to answer specific questions
- Collecting data using appropriate strategies
- Analyzing collected data and summarizing findings
- Using the findings

Evaluation is an active, dynamic process. It is the ability to do evaluation—to assess what is taking place with a program, a strategy, or an initiative—that allows an organization to use data to inform its decision making. For evaluation to be relevant, to reach its potential, it must respond to important questions, be timely, and provide a springboard for

_____
*Note that evaluation in the context of this guidebook is not meant to include **experimental design or research** processes such as randomized controlled trials or multi-year longitudinal research projects.

action. Evaluation is what grounds learning, continuous improvement, and data-driven decision making.

Organizations with evaluation capacity are those with staff who are able:

- To do, commission, or contribute, such as through participatory evaluation, meaningful and timely evaluations
- To understand information collected through evaluation
- To use information from an evaluation to make decisions

Evaluation capacity used well leads to programs, strategies, and initiatives that allow organizations to better deliver on their missions and to better meet the needs of those they serve. ***But because of the focus on service delivery, it does not necessarily lead to more effective organizations.***

Evaluative Thinking is a type of reflective practice that uses the same five evaluation skills listed above in areas other than programs, strategies, and initiatives. It is an approach that fully integrates systematic questioning, data, and action into an organization's work practices. It can contribute to improved learning and decision making.

Evaluative thinking is grounded in the same approach as evaluation: Better planning and decision making happen when they are based on systematically collected data.** (*For example, an organization that applies evaluative thinking to HR strategies*

determines needs and existing levels of competence before procuring training, assesses both the delivery of the training and the use of material that were introduced, and then completes the process by using the information to plan future trainings.)

Evaluative Capacity is the combination of evaluation skills and evaluative thinking. It requires a commitment to doing and using evaluation in programs, strategies, and initiatives *as well as* a commitment to using those same skills in other aspects of organization work.**

Organizations with high evaluative capacity:

• Use the five evaluation skills in programmatic AND organizational work

• Frame questions which need to be asked and can be answered with data

• Have staff who have and hone their evaluation skills

• Have staff who use evaluation skills to collect and analyze data

• Have a clear plan of what to do with the information they collect

• Actually use data for making decisions

• Act in a timely way

Organizational Effectiveness, Evaluative Capacity, and Evaluative Thinking

Organizational effectiveness has been operationalized by the Grantmakers for Effective Organizations (GEO) (www.geofunders.org). They state that it is the ability of an organization to fulfill its mission through a blend of sound management, strong governance, and a persistent rededication to achieving results.

**This guidebook provides examples of and strategies for increasing evaluative thinking in the following areas: mission development, strategy planning, governance, finance, leadership, fund development/fund raising, evaluation, program development, client relationships, communication and marketing, technology acquisition and planning, staff development, human resources, business venture development, alliances and collaborations.

For an organization to be able to determine the soundness of its management, the strength of its governance, and its capacity/status regarding achievements of results, that organization *must* have evaluative capacity, **both** evaluation skills **and** the ability to transfer those skills to organizational competencies.

Evaluative capacity does not automatically bring about organizational effectiveness. But, without evaluative capacity, it is difficult for an organization to be able to understand or assess its level of effectiveness or to determine when strategies designed to enhance organizational effectiveness are in fact doing so.

The remainder of this guidebook clarifies ways that organizations can develop and strengthen their **evaluative capacity** by identifying and enhancing **evaluative thinking** and by building additional **evaluation capacity**.

Indicators of Evaluation in Multiple Organizational Areas

Through intensive work with leaders and staff from nonprofit organizations in Rochester, New York, the Bruner Foundation identified 15 organizational areas where evaluative thinking is particularly important. Multiple indicators for each area are listed in the following charts, and details about how to achieve them are addressed throughout the rest of this guidebook.

To the right are multiple indicators within each organizational capacity area. The list is not exhaustive nor does the absence of an individual indicator imply that evaluative thinking is absent or even substantially reduced at the organization. Rather, this is a starting list of indicators, shown to promote reflection about evaluative thinking in a systematic way. This set of indicators was identified through work with multiple organizations and has been changed, upgraded, and enhanced over time. Users of this guidebook are encouraged to think about these and other ways that evaluative thinking is demonstrated (or absent) in their own organizations.

Indicators of Evaluative Thinking	
Organization Mission	**Organizations that use Evaluative Thinking will:** • Develop mission statements specific enough to provide a basis for goals and objectives. • Review and revise the mission statement on a scheduled basis (e.g., annually) with input from key stakeholders as appropriate. • Regularly assess compatibility between programs and mission. • Act on findings of compatibility assessments (in other words, if a program is not compatible with the mission, it is changed or discontinued).
Strategic Planning	• Have a formal process for strategic planning. • Obtain input from key stakeholders (staff, board, community, and clients) about strategic direction where appropriate using evaluative strategies such as interviews and surveys. • Assess activities related to strategic process at least annually and involve key stakeholders (staff, board, community, and clients) in assessment where appropriate. • Use strategic plans to inform decision making.
Leadership	**Organizational leaders who use Evaluative Thinking will:** • Support and value evaluation and evaluative thinking. Use evaluation findings in decision making. • Include attention to evaluation as an important part of a succession plan. New leaders will be expected to value and be knowledgeable about evaluation. • Educate staff about the value of evaluation and how to participate effectively in evaluation efforts. • Motivate staff to regularly use specific evaluation strategies. • Modify the organizational structure as needed to embrace change in response to evaluation findings. • Foster use of technology to support evaluation and evaluative thinking. • Use data to set staff goals and evaluate staff performance. • Use data to make staffing decisions (e.g., to decide which staff works on which projects, which staff members are eligible for promotions or advancements, or which staff members need more assistance). • Include attention to evaluation in all management-level succession planning. Managers should be expected to value and where possible be knowledgeable about evaluation.
Governance	**In Organizations where there is Evaluative Thinking:** • The board uses appropriate data in defining its goals, work plan, and structure to develop plans summarizing strategic direction. • The board regularly evaluates its progress relative to its own goals, work plan, and structure. • The relationship between the organization mission and plans for strategic direction are assessed. • There is a systematic process and timeline for identifying, recruiting, and electing new board members. • Specific expertise needs are identified and used to guide board member recruitment. • The board regularly (e.g., annually) evaluates the executive director's performance based on the established goals and work plan, including, where appropriate, the use of some program evaluation results. • The board assesses the organization's progress relative to long-term financial plans. • The board assesses the organization's progress relative to evaluation results.

Indicators of Evaluative Thinking	
Finance	• Systems are in place to provide appropriate financial information to staff, and board members are monitored to ensure they inform sound financial decisions. • A comprehensive operating budget which includes costs for all programs, management and fundraising, and identifies sources of all funding is developed and reviewed annually. • Unit costs of programs and services are monitored through the documentation of staff time and direct expenses. • Financial status is assessed regularly (at least quarterly) by board and executive leaders. • Year-end revenues and expenses are periodically forecast to inform sound management decisions. • Financial statements are prepared on a budget versus actual and/or comparative basis to achieve a better understanding of finances. • A review process exists to monitor whether appropriate and accurate financial information is being received from either a contracted service or internal processing. • Capital needs are reviewed annually. • The organization has a plan identifying actions to take in the event of a reduction or loss in funding.
Fund Raising/Fund Development	**Organizations that use Evaluative Thinking will:** • Conduct research on potential fund development opportunities (grants, contracts) and assess which to pursue. • Develop a written fund development plan that clarifies which grants and contracts will be pursued. Assess whether the plan is being followed and why changes and exceptions are made. Revise the plan as needed based on assessments. • Involve program staff in proposal writing, especially sections on program design and outcomes on which the program will be assessed. • Assess the costs and benefits for fund-raising events and activities.
Business Venture Development	• Systematically identify gaps in community service. • Assess whether they have the capacity to bring in new types of business. • Research new business venture opportunities. • Base new venture development on capacity findings, results of gap studies, and business venture development research.
Technology Acquisition & Training	**In Organizations where there is Evaluative Thinking:** • Assessment processes will be in place to make decisions about technology maintenance, upgrades, and acquisition. • Technology systems include software that can be used to manage and analyze evaluation data (e.g., Excel, SPSS). • Technology systems provide data to evaluate client outcomes. • Technology systems provide data to evaluate organizational management. • Technology systems are regularly assessed to see if they support evaluation. • Staff technology needs are regularly assessed.
Client Interaction	• Client needs assessments are conducted regularly, and client services are designed in response to determined needs. • Client satisfaction is regularly assessed and the results of client outcome assessments and client satisfaction are used in development of new programs.

Indicators of Evaluative Thinking	
Marketing & Communications	**Organizations that use Evaluative Thinking will:** • Have a marketing and communications plan. The plan will be linked to the organization's strategic plan and will be used to help the organization achieve its mission. • Involve multiple stakeholders, including staff, board members, and technical assistance providers, as needed, to develop and assess the marketing and communications plan. • Assess the effectiveness of the organization's marketing and communications plan (i.e., determine whether an accurate message is getting out and whether delivery of the message is furthering the mission of the organization.)
Program Development	• Identify gaps in community services before planning new programs. • Incorporate findings from evaluation into the program planning process. • Involve multiple stakeholders in developing and revising program plans. • Develop written program plans including a logical formulation of each program. • Follow program plans where possible and insure that there are strategies in place to modify program plans if needed.
Program Evaluation	• Regularly conduct evaluations that include attention to characteristics, activities, and outcomes of selected programs. • Involve program staff, organization leaders, and clients (as appropriate) in developing and revising program evaluation plans as well as collecting and analyzing program evaluation data. • Share results of program evaluations including findings about client outcomes, as appropriate, with leaders, staff, clients, board members, and funders. • Use results of program evaluation to drive continuous improvement of programs. • Insure that there are key staff with evaluation expertise to address the organization's evaluation needs and that there are staff members whose jobs or components of their jobs are dedicated to evaluation. • Hire evaluation consultants when needed. • Provide or obtain training in evaluation for program staff members and make sure that the training is current, well-delivered, and provided for enough staff members to insure that evaluation use is a standard practice.
Staff Development	• Conduct a formal staff development needs assessment annually. • Develop a plan for staff development based on needs assessment data. • Evaluate the staff development plan. • Provide opportunities for staff to assess staff development training sessions. • Use results of staff training assessments to influence future staff development.
Human Resources	• Have an established personnel performance review process. • Use performance reviews to provide feedback relative to performance expectations. • Collect and update information on credentials, training and cultural competencies of staff; and then use the results to recruit, hire, and train culturally competent staff. • Conduct regular (e.g., annual or biannual) staff satisfaction surveys and use the results to inform modification of policies and procedures.
Alliances & Collaboration	• Evaluate existing partnerships, alliances, and collaborations based on organization mission and strategic plans. • Identify additionally needed partnerships, alliances, and collaborations. • Regularly assess partnerships, alliances, and collaborations in which the organization is involved to determine if they are functioning effectively and continue to meet organization mission and strategic direction.

Why and How to Assess Evaluative Thinking in Your Organization

In 2005, Bruner Foundation evaluation consultants and representatives from 12 nonprofit organizations in Rochester, New York created the *Evaluative Thinking Assessment Tool* to assess the extent to which evaluative thinking is present in various organizational capacity areas. The tool grew out of the Bruner Foundation's Evaluative Thinking in Organizations Study (ETHOS). For more information about ETHOS, please see the *Effectiveness Initiatives* section of the Bruner Foundation website (www.brunerfoundation. org). In 2007, the tool was automated and in 2010–11, it was updated again after additional study.

The modified version of the Evaluative Thinking Assessment Tool was designed to capture leader perceptions about evaluative thinking in a critical subset of organizational capacities, for a particular point in time. Both the original and the modified version of the Evaluative Thinking Assessment Tool include multiple questions to capture information about indicators of evaluative thinking in the 15 different organizational capacity areas. For each item on the modified version of the assessment tool, the respondent is asked to report whether an indicator of evaluative thinking is present using codes shown in drop-down boxes next to each indicator. It is expected that summarizing the organization's best projections about evaluative thinking will help the organization recognize whether and to what extent it is incorporating specific evaluative thinking strategies into its work and in which organizational areas. It will also help the organization to prioritize strategy changes related to evaluative thinking.

Those completing the Bruner Foundation Evaluative Thinking Assessment Tool are advised to complete all 15 worksheets in the instrument and then view the summary table and summary graph that are generated automatically. Summaries and all pages in the Evaluative Thinking Assessment workbook can be printed if desired; users of the tool are encouraged to generate their own Evaluative Thinking Assessment report from these worksheets. Please see the actual Evaluative Thinking Assessment tool for remaining detailed instructions for its use, as well as the example report which presents EvaluativeThinking Assessment results for a fictitious organization. [These tools can be found in the Effectiveness Initiatives section of the Bruner Foundation website: *Featured Resources* on the Evaluative Thinking page.]

The Evaluative Thinking Assessment Tool was designed to facilitate discussions about:

- Perceptions of evaluative thinking in multiple organizational areas
- Changes in evaluative thinking
- Challenge areas where additional evaluative thinking might be incorporated into organizational work

Evaluative Thinking Assessment scores can also inform the setting of priorities regarding incorporation of or enhancement of evaluative thinking in organizational practice. Users of the tool are also encouraged to think about score thresholds for their own organizations—what is ideal, what is expected, and what is unacceptable—and to think of responses to challenge areas that are identified through its use.

Evaluative Capacity and Organizational Leaders

In order for an organization to develop and sustain evaluative capacity, it must have leaders who, at the very least, not only value evaluation, but more importantly, practice evaluative thinking.

Specifically, organizational leaders should:

- Support evaluation and evaluative thinking
- Provide or obtain training in evaluation and evaluative thinking for themselves and key staff
- Use data to set staff goals, evaluate staff performance, and make staffing decisions
- Include attention to evaluation as an important part of succession planning
- Foster the use of technology to support evaluation and evaluative thinking
- Use evaluation findings in decision making

Leaders and key staff should also have regular involvement in evaluation actions,
including some or all of the ways identified in the list below.

Evaluative Roles of Organization Leaders and Key Staff

- Identify programs that will benefit from evaluation
- Identify and support staff in all aspects of their efforts to design and conduct evaluations
- Promote the widespread practice of assessment and learning within the organization
- Identify questions of consequence (based on expected outcomes)
- Carefully select evaluation strategies that will produce usable evaluation findings
- Thoughtfully conduct evaluation activities and analyses of data
- Use evaluation results to strengthen programs and enhance program decision making
- Share evaluation training and other evaluation learning with colleagues
- Promote clear communication about the purposes of the evaluation
- Design, conduct, and support evaluations that are honest and that help inform decision making
- Promote the continuous improvement of programs, as needed
- Set short-term measures and milestones, but also identify longer-term opportunities for organizations to use information and produce stronger programs and desired outcomes
- Identify outcome targets in advance of initiating evaluation (use prior efforts, external standards, or agreed-upon expectations to determine reasonable targets)
- Do peer education with colleagues
- Support policies that clarify:
  —How evaluation results should be used (i.e., in context, and in conjunction with other findings and documentation)
  —Who is responsible for evaluation (including quality assurance, training, compensation)
  —How evaluative capacity can be sustained (e.g., through succession planning that includes attention to evaluation knowledge, through technology planning that supports evaluation)

As stated in the acknowledgements to this guide-book, it is the belief of the authors and their many partners over time that building a culture of evaluation—one in which evaluation skills are firmly embedded and used for development, improvement, and decision making about programs, strategies, and initiatives—is fundamentally important in creating a learning organization, a data-driven organization, an organization sensitive to making informed decisions in a timely way. We also believe that evaluation skills are highly transferrable beyond evaluating an organization's programs, strategies, and initiatives. Evaluative thinking is just that—a way of thinking, a process of acquiring and using relevant data in any aspect of an organization's work. The following section provides additional examples of the application of evaluative thinking to important organizational competency areas **beyond** program evaluation.

2. Evaluative Thinking at Work

As stated in the Introduction, evaluative thinking can be used in many areas of organizational work. Here we address technology, staff development and HR, and collaboration. In any area, the process is the same: ask important questions, systematically collect and analyze data, summarize and share information, and act on the findings.

Evaluation and Technology

Organizations typically need technology systems for four main purposes: (1) management support (e.g., for use with financial, HR, or procedural information); (2) internal and external communication (e.g., e-mail capability, proposal or report development and submission, contact with program participants and others); (3) program operations (e.g., client tracking, archiving and recordkeeping, best practice research, program descriptions, logic model development); and (4) evaluation (for example, of programs, clients, overall organizational effectiveness).

An organization that is committed to evaluative thinking uses it to inform decisions related to technology and embraces supportive technology use. In other words, they ask questions about whether the organization's hardware and software truly inform organizational and program practices, they systematically collect data about technology systems in use to ensure that they truly are supportive, and they analyze and act on the data that is collected (i.e., technology changes are made based on information). In

addition, they use technology to support evaluation work. The list on the next page enumerates questions that should be answered to determine if an organization is using evaluative thinking to support technology and using technology to support evaluation.

Questions to Consider Regarding Use of Evaluative Thinking to Determine Technology Needs and Use

- Do staff who need them have individual computers or access to computers? Do most or all staff members use the Internet as a research tool? Do most or all staff members use e-mail for internal communication?

- Does the organization have technical support for computers and software (e.g., assistance for computer problems)? Is there a technology systems plan and budget in place to guide decisions about technology, maintenance, upgrades, and acquisitions?

- Are the technology needs of staff and/or programs regularly assessed?

- Are the organization's technology systems regularly assessed to see if they support organizational effectiveness and evaluation in particular?

- Are the organization's technology systems changed or upgraded if assessment data suggest there are unmet needs?

- Do staff members receive training to use technology to track clients and support program delivery and evaluation?

Questions to Consider Regarding Use of Evaluative Thinking to Determine Technology Needs and Use

- Do staff who need them have individual computers or access to computers? Do most or all staff members use the Internet as a research tool? Do most or all staff members use e-mail for internal communication?

- Does the organization have technical support for computers and software (e.g., assistance for computer problems)? Is there a technology systems plan and budget in place to guide decisions about technology, maintenance, upgrades, and acquisitions?

- Are the technology needs of staff and/or programs regularly assessed?

- Are the organization's technology systems regularly assessed to see if they support organizational effectiveness and evaluation in particular?

- Are the organization's technology systems changed or upgraded if assessment data suggest there are unmet needs?

- Do staff members receive training to use technology to track clients and support program delivery and evaluation?

- Do staff members have opportunities to influence design of or use of technology for tracking information about their programs?

- If the organization is using technology to support evaluation, do the organization's technology systems:
  —Store and manage data to evaluate organizational effectiveness overall including financial, governance, communications, and staffing information, as well as data used to evaluate client outcomes?
  —Include software (such as SPSS or SAS) that can be used to manage and analyze data?

- Do staff members have opportunities to influence design of or use of technology for tracking information about their programs?

- If the organization is using technology to support evaluation, do the organization's technology systems:
  — Store and manage data to evaluate organizational effectiveness overall including financial, governance, communications, and staffing information, as well as data used to evaluate client outcomes?
  — Include software (such as SPSS or SAS) that can be used to manage and analyze data?

When Does an Organization Need Technology Help to Support Evaluation and Evaluative Thinking?

Most program evaluation and evaluative thinking technology needs are straightforward—the technology must only help to manage small databases and

to perform simple calculations such as frequencies (i.e., counts of occurrences or responses), and cross-tabulations (analyses of combined data). Staff can use existing technology (hardware, software, internet connections) to manage and store information and to conduct analyses of data to inform decision making.

There are at least four conditions, however, when additional support is beneficial:

- When there are no staff members with expertise in the use of software to manage and analyze data;
- When the evaluations, plans or assessments being conducted require the development of, or access to, large or complex databases;
- When a funder or collaborative agreement requires that information be stored in and analyzed from customized databases;
- When the planned analyses cannot be conducted by hand or with the use of readily available software such as Excel.

It is very important to factor in evaluation needs when considering new technological support. *In organizations where there is evaluative thinking, technology planning processes include attention to evaluation and ways that technology can support it.*

Using Your Existing Technology—What Else Is Out There?

Most organizations have multiple computer workstations for staff member use. These computers are commonly equipped with the Microsoft Office suite of software that includes word processing, spreadsheet (Excel), and relational database (Access) applications, with connections to the Internet. (Less common are staff members who are able to use all of these applications specifically to support program evaluation and other evaluative thinking needs.)

Excel, in particular, is a versatile and readily available tool for managing and analyzing both survey and record review (administrative) data. For simple instructions and a sample Excel file including multiple worksheets, visit the resources and tools' pages of the Bruner Foundation website.

There are four other types of software that can address data management and analysis needs to help support regular use of evaluative thinking and effective evaluation. Consider whether any of these are necessary for your continued efforts to use technology evaluatively:

1. Statistical analysis software (such as SPSS, Minitab or SAS)
2. Software to help conduct electronic/web-based surveys (such as Survey Monkey or Zoomerang)
3. Qualitative data analysis software (such as QSR N6, Atlas.ti, E6, QSR NVivo)
4. Database software (such as FileMaker Pro, QuickBase)

Again, in terms of organizational priorities, it is important to both use evaluative thinking to enhance technology and to use available technology systems to enhance your evaluation.

How Do Nonprofit Organizations Get Technology Help?

1. **Use or develop internal capacity.** Be sure to develop clear job descriptions, and/or needs statements, and to assess the effectiveness of technology staff.

2. **Seek technical assistance from the nonprofit or for-profit sectors** in the form of interns, consultants, or shared services. Be sure to assess the usefulness of the services.

3. **Seek capacity-building grants from funders** offered exclusively or as part of a larger effort. Be sure the grant's intent matches the organization's technology needs.

Evaluative Thinking, Staff Development, and HR

In addition to the various data collected about programs and participants, most organizations manage data about staff. This includes information about staff background and qualifications, compensation, and performance. Many organizations also provide ongoing training for their staff, both internally and via other providers, including some professional development that is mandatory and some that is elective. Evaluative thinking can be useful in both of these areas of organizational function.

Organizations that incorporate Evaluative Thinking into their Human Resources practices:

- Have an established personnel performance review process
- Use performance reviews to provide feedback relative to performance expectations
- Collect and update information on credentials, training and cultural competencies of staff and use the results to recruit, hire, and train culturally-competent staff
- Conduct regular (e.g., annual or biannual) staff satisfaction surveys and use the results to inform modification of policies and procedures

Organizations that incorporate Evaluative Thinking into their Staff Development practices:

- Conduct a formal staff development needs assessment annually

- Develop a plan for staff development, based on needs assessment data

- Evaluate the staff development plan

- Provide opportunities for staff to assess staff development training sessions

- Use results of staff training assessments to influence future staff development

Organizations can learn from and about their staff using evaluation data collection strategies such as surveys, interviews, and record reviews. As shown in the table on the next page, information about staff and their actions, perceptions, and feedback can be used to inform organizational practices. Remember: *Organizations that regularly use evaluative thinking ask questions of substance (including questions about their staff), determine what data are needed to address the questions and how those data could be obtained, systematically collect and analyze data (including staff feedback), and share the results and develop strategies to act on findings.*

Staff Data to Collect—Examples	Surveys	Interviews	Record Reviews	How Staff Data Can be Used
Staff Satisfaction Data (including assessment and understanding of or agreement with organizational mission, policies, and procedures)	✓✓	✓		Inform efforts to promote staff retention; revise, continue, or introduce policies or procedures based on feedback Determine alignment with organizational mission Investigate work schedule options and incentives
Staff Background Information* (including skills, experiences, interests, training, educational achievements, and educational levels) * Note that these must be updated regularly to be useful.	✓✓	✓	✓✓	Determine composition of the staff, identify specific staffing needs, update education, training and experience information Determine suitability of placement, supervision Identify staff resources (e.g., capability to provide training to others)
Job-Related Behaviors and/or Preferences (including reports on how and how often certain job tasks are undertaken, and/or how much time is needed for them)	✓	✓		Revise, continue, or introduce policies and procedures Determine alignment with organizational philosophies, policies, and procedures Determine education or communication needs
Staff Development Feedback	✓✓	✓		Determine staff development needs for different positions Obtain feedback regarding response to and usefulness of staff development training
Benefits and Compensation Information	✓	✓	✓	Obtain information about how staff understand and use benefits (e.g., so that benefits packages can be assessed and modified as needed) Assess and compare compensation levels by various partitions such as experience, length of employment, etc.

✓✓ = preferred method   ✓ = acceptable methods

Evaluative Thinking and Collaboration

It is often advantageous (sometimes even required) for organizations to work together. The combined efforts of two or more organizations are identified by a variety of terms with overlapping definitions. Common forms of working together include:

- **Alliances:** Formal agreements establishing an association between organizations to achieve a particular aim or advance common interests.

- **Partnerships:** Cooperative, usually voluntary, relationships between groups that agree to share responsibility for achieving some specific goal(s). Partnerships can also include the sharing of resources, risks, and rewards according to terms reached through (often prespecified, sometimes legal) agreement.

- **Collaboration:** Though it is often used as a synonym for the above and to describe other joint or multiple efforts, this term is primarily meant to signify durable and extensive relationships where participants bring separate organizations into a new structure with full commitment to a common mission. Partners pool or jointly secure resources and share the results and rewards.

Regardless of what the relationship is called, working together requires decision making. Decision making is strengthened by evaluative thinking.

Should you join in or go it alone?

When forming alliances, partnerships, or collaborations, always consider the following.

- Alliances, partnerships, and collaborations are not efficient ways of doing business.

- All partners must do **and** get something, but their efforts are not necessarily equal.

- True collaboration is inherently interactive, relationship-based, and voluntary.

- It takes time, is difficult to achieve, and cannot be mandated, forced, or faked.

To apply evaluative thinking to consideration of partnership formation, try collecting and analyzing data to answer the following questions.

- Why is a combined approach better? Are there additional or better outcomes that will result from the combined approach?

- Are partners necessary to achieve the desired outcomes?

- Which organizations and which specific staff or positions from each organization will be involved and how?

- What is each organization's history of and commitment to working with others?

- How will working together work? Are there written, agreed upon guidelines and responsibilities for key collaborative functions such as oversight, communication, recordkeeping, and decision making?

- What is the timeframe for collaborative work? How much time will be required for start-up, maintenance, and dissolution of the collaboration?

- How will the effectiveness of work done with others be evaluated? Is the group committed to using evaluation findings to strengthen its work?

Once partnerships have formed, evaluative thinking can be used again to determine the effectiveness of partnerships and collaborations and inform course corrections.

- Are partnerships, alliances, and collaborations embraced by the organization?

- Do partners articulate a shared vision for their work? Can they articulate how they benefit from participating in the collaboration? Do partners support the partnership agenda?

- Are the short- and long-term goals, objectives, and strategies of the partnership clear to all partners? Are they realistic? Are goals, objectives, and strategies revised with new data regularly?

- Are existing partnerships, alliances, and collaborations evaluated by the organization?

- Can partners articulate how the advantages of membership offset direct and other partnership costs? Is there evidence of appreciation and respect for different stakeholder roles and organizational differences?

- Are partnerships, alliances, and collaborations functioning effectively?

  — Do the partnerships have the appropriate and required mix of participants even after there is turnover or membership change? Is there targeted recruitment?

  — Are the partnerships able to handle difficult discussion and resolve conflict respectfully and consensually?

  — Do formal and informal communication mechanisms exist? Are they used regularly?

  — Do all partners clearly understand their roles and responsibilities? Are memorandums of understanding (MoUs) or other formal agreements and accountability mechanisms in place regarding how the partnership works?

  — Are decision making and direction setting shared by all active members of the partnership?

  — Are meetings regular, well-organized, well-run? Is attendance consistent? Are there effective committees and strong leaders for the partnership? Do partners formally interact regularly and often? Do they update one another, discuss issues openly, and convey all necessary information to one another and to people outside their group? Do meetings include time for formal discussions? Are most members engaged in meetings and the work of the partnership?

- Do partnerships, alliances, and collaborations continue to meet the organizational mission and strategic direction?

- Is there a plan to identify additional needed partnerships, alliances, and collaborations?

Results of evaluative inquiries (i.e., answers to the questions posed in the preceding section) can and should be used to help organizations determine if new collaborations are needed and/or are an appropriate fit for the organization. Results of inquiries about existing partnerships (i.e., answers to the questions shown above) can be used to inform decisions about continued or changed collaborative involvement.

Using Evaluative Thinking to Improve Communications and Marketing

Organizations report a great deal of information about their services and operations via websites, blogs, annual reports, evaluation reports, and other publications or sites. One important indicator of evaluative thinking related to communications and marketing is the development of a communications and marketing plan. When organizations regularly use evaluative thinking, the following will be evident:

- A communications and marketing plan linked to the organization's strategic plan is used to help the organization achieve its overall mission and objectives;

- The communications and marketing plan is developed with input from multiple stakeholders including staff and board members and external technical assistance providers as appropriate;

- The effectiveness of the organization's communications and marketing plan (i.e., whether an accurate message is getting out and whether delivery of that message is furthering the mission of the organization) is regularly assessed.

Further, communications and marketing plans will include and be linked to evaluation results. _In organizations that regularly use evaluative thinking, the content of what is marketed and included in communication is driven, in part, by evaluation results._

3. Evaluation Essentials

As stated in Section I, it is the belief of the authors and their many partners over time that building a culture of evaluation is fundamentally important for data-driven learning organizations. Knowledge about essential evaluation skills is key. Further, it is our belief that evaluation skills are highly transfer- rable beyond evaluating an organization's programs, strategies, and initiatives and can be applied to any aspect of an organization's work. This section provides an important overview regarding essential evaluation skills.

Evaluation is a science with a relatively short his- tory. It became a distinctive field of professional social science practice in the late 1960s (Patton 1982) and is widely used in conjunction with the work of nonprofit organizations today. Many types and classifications of evaluation exist; many terms are associated with the practice. (For a brief list of important evaluation terms and a brief summary of the history of evaluation, see Appendices 1 and 3.) Because evaluation is the cornerstone to the development of evaluative capacity, this guidebook includes two sections (this and the following sec- tion) detailing basic content about evaluation and strategies for enhancing evaluation capacity.

It is important to start any discussion of evaluative capacity with two key definitions.

Working Definition of Program Evaluation: The practice of evaluation involves the thoughtful, systematic collection and analysis of information about the activities, characteristics, and outcomes of programs for use by specific people to reduce uncertainties, improve effectiveness, and inform decisions regarding those programs (adapted from Patton 1982).

Working Definition of Evaluative Thinking: A type of reflective practice that incorporates use of sys- tematically collected data to inform organizational actions and other decisions.

Evaluation capacity is the ability of staff and their organizations to do evaluation. Evaluative capacity is the combination of evaluation skills and evalua- tive thinking. It requires a commitment to doing and using evaluation in programs, strategies, and initia- tives *as well as* a commitment to using those same skills in various aspects of the organization.

For more content, activities, and a step-by-step guide to plan for and conduct program evaluation or to commission external evaluation, see Participatory Evaluation Essentials: *An Updated Guide For Non- profit Organizations And Their Evaluation Partners* (Baker 2010). This guidebook can be downloaded at no cost from the Effectiveness Initiatives section of the Bruner Foundation website, www.brunerfoun- dation.org.

Evaluation Planning

Planning an evaluation requires forethought and attention to strategic and technical steps. It also requires recognition of assumptions about the work being evaluated and existing biases held by those who will conduct the evaluation. Good evaluation planning also includes front-end analysis planning and thought about use of evaluation results. There is a need to establish neutrality to the degree possible by those making evaluation decisions. A written plan or design is required for the evaluation overall. Spe-cific administration and analysis plans are strongly recommended for each data collection strategy.

Evaluation Strategy Clarifications

In addition to basic definitions, it is useful to under-stand current and common thinking about evaluation strategy. To that end, it is important to recognize the following:

- All evaluations are **partly social**, because they involve human beings; partly political, because knowledge is power and decisions about what gets asked and how information is used are political decisions at the organizational level; **partly techni-cal**, because choices always have to be made about how to collect and analyze data (Herman, Morris, Fitz-Gibbons 1996).

- Both **qualitative** (descriptive) and **quantitative** (numerical) data can be collected using multiple methods (e.g., observations, interviews, surveys, or statistical analyses of practical assessments from review of program records). Although there has been much debate about which strategies and types of data are best, current thinking indicates that **both qualitative and quantitative data are valuable and both can be collected and analyzed rigorously**.

- There are multiple ways to address most evaluation needs. While there are things you should definitely avoid, there is **no single or best way to design and conduct evaluation**. Different evaluation needs and resources call for different designs, types of data, and data collection strategies. A rigorous evalua-tion will include:

  —Multiple data collection strategies (e.g., some combination of surveys, interviews, observations, and record reviews, though not necessarily all in one evaluation)

  —Collection of data from or about people from multiple perspectives (e.g., program partici-pants, their caregivers, staff of the program, or other service providers)

  —Collection of data at multiple points in time

In brief, evaluation involves five steps.

1. Specifying questions

2. Developing an evaluation design

3. Applying evaluation logic (i.e., identifying out-comes, indicators and targets, and determining how and when to measure them)

4. Collecting and analyzing data, making action recommendations

5. Summarizing and sharing findings

Purposes of Evaluation

Evaluations are typically conducted to accomplish one or more of the following:

• Render judgments

• Inform data-driven decision making

• Facilitate improvements

• Generate knowledge

It is critical to carefully specify why evaluation is being conducted. This should be done at the earliest stages of evaluation planning and with the input of multiple stakeholders (e.g., organization leaders, staff, and possibly board members, program participants, and funders).

The Importance of Evaluation Stakeholders

Evaluation stakeholders are people who have a stake—a vested interest—in evaluation findings (Patton 1997, p. 41). Stakeholders include anyone who makes decisions about a program, desires information about a program, and/or is involved directly or indirectly with a program.

• Most programs have multiple stakeholders.

• Stakeholders typically have diverse and often competing interests (e.g., if you were to ask what was important about an afterschool program of youth participants, parents, teachers, and program staff, the answers and related program interests and needs would be different).

• In program evaluation, stakeholders typically include organization officials (e.g., executive director), program staff, program clients or their caregivers, and ultimately program funders. Sometimes

organization board members, community members, or other organizations are also stakeholders. **It is critical for an evaluation to involve some of these key stakeholders in the evaluative process, especially in evaluation planning**. When appropriate, there can also be important roles for stakeholders in data collection, analysis, and reporting. *When stakeholders are not involved at all in evaluation, there will likely be misunderstandings and perhaps dissatisfaction regarding evaluation design, implementation, and use*.

Evaluation Design

Developing an evaluation design helps you think about and structure an evaluation. Evaluation designs communicate evaluation plans to evaluators, program officials, and other stakeholders.

Evaluation is not free! But evaluation should not be viewed as in competition with program resources.

Evaluations can be funded as a component of a program (using program funds) or as a separate project (using ear-marked or additional funds).

A common rule of thumb is to set aside at least 10% of the cost of the program for evaluation. For additional details about evaluation funding, see Appendix 9.

A good design should include the following:

• Summary information about the selected program, why it is to be evaluated, what the underlying logic is regarding the program (i.e., how the program is expected to use resources to deliver services that contribute or lead to outcomes, and how it will be known when outcomes have occurred)

• Between two and five questions to be addressed by the evaluation

Integrating Evaluative Capacity Into Organizational Practice • Section 3

- Clear description of data collection and data analysis strategies that will be used
- Identification of individuals who will be responsible for collecting data
- Estimated timeline for evaluation activities and projected costs to do the evaluation
- Description of products of the evaluation (usually including report and executive summary with formatting/layout information) with notation regarding who will receive them and probable plans for their use

Evaluation Questions Get You Started

Evaluation questions ideally should be determined by the service providers (e.g., executive directors, program directors, and managers, maybe line staff and other stakeholders), together with the evaluator, in accordance with the purpose of the evaluation.

Evaluation questions should:

- Focus and drive the evaluation. They should clarify what will and will not be evaluated.
- Be carefully specified (and agreed upon) in advance of other evaluation design work.
- Generally represent a critical subset of information that is desired to address the purpose of the evaluation. (While it's easy to identify many questions that could and maybe should be answered, usually only a subset can be addressed in any single evaluation due to resource constraints and the need to focus attention on a manageable set of requests).

For more information and examples of evaluation questions as they relate to different evaluation purposes, see Appendix 4.

Criteria of Good Evaluation Questions

It is important to keep the number of evaluation questions manageable. The exact number depends on the purpose of the evaluation and resources available to conduct the evaluation; however, limiting the

evaluation to address between two and five questions is strongly advised.

The following are criteria of good evaluation questions (adapted from Patton 1997):

- It is possible to obtain data to address the questions. The data must be available for use by those undertaking evaluation.
- Questions should focus on current or recent (but not future) events, skill development, or response to programs.
- There is more than one possible answer to the question, i.e., the findings are not predetermined by the phrasing of the question.
- Those conducting the evaluation want and need the information derived from addressing the questions and know how it will be used internally and, where appropriate, externally.
- The questions are aimed at changeable aspects of programmatic activity (i.e., they should focus on those things which can be modified where findings warrant change).

> **Evaluation does not focus on cause or attribution nor is it conducted to prove anything. That requires more of a research approach, including experimental design (see Appendix 5). Answering evaluation questions that meet the above criteria allows for data-driven decision making about important, select programs or program features.**

Additional Evaluation Design Considerations

As evaluation is planned, the following should be determined:

1. Is assistance from a professional evaluator necessary, useful, or affordable? (See the following section on commissioning evaluation.)

20   A Bruner Foundation Guidebook for Nonprofit Organizations and Their Stakeholders • Copyright © by the Bruner Foundation 2012

2. What descriptive information should be collected to clarify context (e.g., staffing plans, activity schedules, history)?

3. How will service delivery and program implementation be documented? What must be tracked (e.g., intake, attendance, activity schedules)?

4. What specific outcomes will be assessed? What are indicators of those outcomes? What targets or levels of outcome attainment are desired? *How good is good enough?*

Different terms are used to describe the results and outcomes of programs, what is expected, and how you know if meaningful outcomes are achieved. The following lexicon, adapted from the United Way of America,

Outcomes are changes in behavior, skills, knowledge, attitudes, condition, or status. Outcomes are related to the core business of the program, are realistic and attainable, within the program's sphere of influence, and appropriate. Outcomes are what a program is held accountable for.*

Indicators are specific characteristics or changes that represent achievement of an outcome. Indicators are directly related to the outcome and help define it. They can be seen, heard or read. Indicators must be measureable and make sense in relation to the outcome whose achievement they signal.

Targets specify the amount or level of outcome attainment that is expected, hoped for, or required. Targets or levels of outcome attainment can be determined relative to external standards (when they are available) or internal agreement (based on best professional hunches, past performance, or similar programs). ***Targets for outcome attainment*** should be established in advance of data collection and analysis whenever possible.

* Note: Though outcomes are commonly described as *changes*, sometimes outcomes are the maintenance of a desired or precondition status.

is useful (though you are reminded that these terms, like all others in evaluation, are used differently by different program stakeholders and evaluators).

Data Collection Overview

There are four primary evaluation data collection strategies which can (and should) be combined to address evaluation questions and allow for multiple sources of data. All have both benefits and limitations and require preparation on the front end. All can be used to collect either qualitative or quantitative data.

• Record and document review

• Surveys

• Interviews

• Observation

Remember, mixed methodologies and multiple sources of data or respondents, collected at multiple points in time, increase evaluation rigor and usefulness of findings. Data do not have to be collected for all participants in every program cycle; determining findings from samples (subgroups) of participants or point-in-time (snapshot) estimates can be useful. Section IV of this guide provides many additional details on data collection.

What Happens After Data Are Collected?

Once evaluation data have been collected via surveys, interviews, observations and/or record reviews, there are three remaining important stages of evaluation.

1. **Data must be analyzed and results must be summarized**. Organizations are great at collecting data, but they often forget to analyze and review the results. Analysis must be done to obtain results and to inform learning and decision making.

2. **Findings must be converted into a format (e.g., a report) that can be shared with others**. The report

may take different forms for different stakeholder audiences (e.g., memo with key findings, oral presentation). The most common form used to summarize evaluation results is the Evaluation Report (see the following and section VI).

3. **Action steps must be developed from findings**. This is probably the most important step—it moves evaluation from perfunctory compliance into the realm of usefulness.

At the organization or program level, making changes based on findings would be next (step 4), often followed by another round of question specification, design development, collection and analysis of data, conversion of findings into a sharable format, and development of new action steps. Both evaluation and evaluative thinking follow these same steps (specify questions, develop designs, collect and analyze data, identify action steps and recommendations, summarize and share findings).

Evaluation Reports

As stated previously, in order to facilitate usefulness, the results of evaluation (or an evaluative inquiry) must be shared. Evaluation reports are the most common strategy for sharing evaluation results. Strong reports are written using straightforward terms in a style mutually agreed upon by evaluators and stakeholders. They typically include:

• A description of the subject program

• A clear statement about evaluation questions and the purpose of the evaluation

• A description of actual data collection methods

• A summary of key findings and discussion or explanation of the meaning and importance of key findings

• Suggested action steps for the program and next steps or issues for further consideration which clarify continued evaluation or program related follow-up (as appropriate)

It is important to note that while the report should conclude with suggested action steps, it is **not** the job of the evaluator to make specific recommendations based on findings (although specific recommendations regarding future evaluation are appropriate). Evaluation stakeholders, i.e., those commissioning the evaluation and those responsible for operations, must determine what actions should be proposed or taken in response to evaluation findings. Their decisions can and should be included in an evaluation report and duly attributed to them. Section VI contains more details about evaluation reporting.

Commissioning Evaluation

Sometimes consultants are needed to conduct an evaluation or to provide coaching and assistance to an organization conducting a participatory evaluation.*** Tips for commissioning and paying for evaluation are provided in Appendices 8 and 9 to this guidebook. Most importantly, organizations that commission evaluation should be sure that the consultants they work with have:

• Basic knowledge of the substantive area being evaluated

• Knowledge about and experience with program evaluation (especially with nonprofit organizations)

• Good references from trusted sources

• A personal style and approach that fits with their organization

---

*** **Working Definition of Participatory Evaluation**: Trained evaluation personnel and practice-based decision makers (i.e., organizational members with program responsibility—service providers) working in partnership. Participatory evaluation brings together seasoned evaluators with seasoned program staff to provide training and then to design, conduct, and use results of evaluation (Cousins 1998).

Things to Determine Before Initiating Evaluation Design

1. **Program and Evaluation Purpose**: What is the general purpose of the selected program, and how does it contribute to the organization's mission? What are the specific reasons why this program was selected for evaluation? How does this project contribute to the broader field? What is the purpose for this evaluation, and what are the possible lessons learned?

2. **Implementation and Feasibility**: Does the target population of the selected program know about and want to participate in the program? If applicable, have all necessary collaborative agreements been secured? How will the organization guard against implementation impediments such as insufficient recruitment or participation, staff turnover, or insufficient program funds?

3. **Program Design and Staging**: What are the key components of the program, how do they fit together, and how are they expected to contribute to participant outcomes? Has a reasonable program logic model been developed? How will an evaluation project best be staged over time?

4. **Outcomes**: What program and participant outcomes are expected? How does the organization know when these outcomes are achieved? If applicable, how are program and participant data tracked? What other reporting is currently being conducted?

5. **External Assistance and Finances**: Will an evaluation consultant be needed to complete the evaluation? If so, how will the consultant be selected? How will the consultant's work be overseen? Will external funding be needed and available to support evaluation of this program? If yes, how can that funding be acquired, and how does the proposed evaluation fit into the mission of the proposed funder?

Note: You don't have to know all the answers to these questions to design an evaluation, but gathering as much information as possible before initiating the design is useful.

Things to Avoid When Working with Evaluation Consultants

- Agreeing to an evaluation design that you do not understand.
- Agreeing to an evaluation where fee payments are not attached to deliverables.
- Commissioning an evaluation on a timetable that is inappropriate for the program.
- Commissioning an overly complicated evaluation design or one for which there is insufficient stakeholder involvement.
- Allowing evaluation consultants to conduct overly-complicated or onerous data collection strategies or to exclude the use of data collected directly from respondents. *At least some primary or direct data collection is preferable*.
- Assuming that an evaluation design must ALWAYS include measurement of outcomes. *Sometimes implementation assessment (i.e. determining if program activities are being delivered effectively) is all that can or should be done.*
- Forcing evaluation of outcomes that are inappropriate or beyond the scope of the program.
- Developing an unrealistic timetable for evaluation reporting. There will always be a lag between final data collection and reporting.

4. Building Evaluation Capacity

As clarified from Section I, the key components of evaluative thinking—**asking questions, systematically gathering data, analyzing data and sharing results, and developing action steps from the results**—can be applied to most aspects of organizational practice. Because we believe that increased capacity to conduct evaluation helps organizations think and act evaluatively, for example when organization administrators and board members use evaluation skills to define, collect data about and determine "compatibility" of programs to mission and then change or discontinue those identified as incompatible, they are using *evaluative thinking* to inform organizational direction. This section of the guidebook provides additional detail for strengthening evaluation capacity, with special attention to effectively collecting and analyzing data.

Data Collection Details

There are four primary ways to collect evaluation data. All have benefits and limitations and many decisions which must be made before data collection can commence. The following are basic distinctions.

Surveys have a series of questions (items) with predetermined response choices. They can include only independent questions or groups of related questions (scales) that can be summarized. Surveys can also include open-ended items for write-in or clarification. Surveys can be completed by respondents or the person(s) administering the survey. They can be administered in write-in, mail-in, electronic, and other alternative formats (e.g., as responses that are cast by holding up a certain colored card, applying colored stickers that signify response categories onto easel sheets with questions, or using candy or marbles of different colors to represent responses).

Interviews are one-sided conversations between an interviewer and a respondent. They can be conducted in person, by phone, with individuals or with groups. Questions are (mostly) predetermined but open-ended. Respondents are expected to answer using their own terms.

Observations are conducted in person to view and hear actual program activities. They can be focused on programs overall or participants in programs.

Record review is a catch-all category that involves accessing existing information. Record review data are obtained from program records including those designed for evaluation, those used for other purposes, and those used by other agencies (e.g., report card grades might be a source of data for evaluation of an after-school program; data collected as part of a drug screening at a health clinic might be used as part of the evaluation of a prevention program).

The table on the following page presents a summary of how different data collection strategies can be used when conducting evaluations. It also provides examples of when various strategies could be used to collect data to inform evaluative thinking for the organization overall (shown in blue).

Use Record Reviews to . . .	**Use Observations to . . .**
Collect behavioral reports.	Document program implementation.
Test knowledge.	Witness levels of skill or ability, program practices, behaviors.
Verify self-reported data.	Determine changes over time.
Determine changes over time.	**Determine whether program modification strategies are being followed, and are helping to respond to challenges.**
Monitor unit costs for a program through documentation of staff time and direct program expense.	
Use Surveys to . . .	**Use Interviews to . . .**
Study attitudes and perceptions.	Study attitudes and perceptions using respondent's own language.
Collect self-reported assessment of changes in response to program.	Collect self-reported assessment of changes in response to program.
Collect program assessments.	Collect program assessments.
Collect some behavioral reports.	Document program implementation.
Test knowledge.	Determine changes over time.
Determine changes over time.	**Determine whether existing technology systems are supporting evaluation needs.**
Ascertain whether staff members are satisfied with their new professional development program, or their medical and dental benefits, and/or still have unmet needs.	

For more details about these data collection methods and practice activities, be sure to review the *Participatory Evaluation Essentials Guide* (Baker and Bruner 2010).

Tips for Improving Data Collection

Maximizing Survey Response

The number one way to maximize survey response is to write or use a good survey instrument and tailor the administration to the desired respondent group. The right set of well-written questions, answer choice appropriateness and brevity are all important. The following are additional tips to increase the number of people who respond.

- Advertise survey purpose and importance as well as administration details in advance.

- Unless anonymous, carefully document who receives and completes surveys. Conduct follow-up and send reminders.

- Consider using incentives including lotteries drawn from respondent lists to encourage responses, but be sure to think carefully about what would "incentivize" your target response group.

- Make it easy for a respondent to answer (e.g., use program time when possible, include postage, be sure return directions are clear, use electronic survey options, send reminder emails, etc).

Nonresponse bias, which happens when large numbers of targeted respondents do not answer, can severely limit your use of response data. If only some of those who should answer a survey actually do, you must limit survey findings only to the group of respondents (e.g., *80% of the survey respondents rated* the activity favorably*). When there is sufficient and representative response, there can be generalization to larger groups (e.g., *80% of participants rated the activity favorably*). See Appendix 10 for an explanation of how to sample to get sufficient response and how many responses are needed to generalize.

Survey Questionnaire Development: Step-by-Step

1. Identify the key issues you wish to explore or 'test' via the survey. Review available literature, including proprietary sources, to determine if there are good surveys or items that already exist to measure the key issues. Start with any known research organizations or projects in the field of interest (e.g., the Harvard Family Research Project [HFRP]).

2. Convert the key issues into questions and remember to:

   • **State the question in very specific terms, using language that is appropriate for the target population**

   • **Use multiple questions to sufficiently cover the topic**

   • **Avoid double negatives**

   • **Avoid asking multiple questions in one**

   • **Be sure response categories match the question, are exhaustive, and don't overlap**

3. Determine what other data are needed for analysis (demographics, other background like how long the respondent has been affiliated with the agency or program, contact information).

4. Determine how the questions will be ordered and formatted.

5. Include directions (such as *mark one for each, mark one or clarify here*) for responses.

6. Have the survey instrument reviewed by others including representatives from the targeted respondent group.

7. Develop an analysis plan and administration plan. That way, you will be prepared for data entry and to summarize findings. Thinking through the analysis plan will also help with survey review—if you can't figure out how you will use data from the survey, you may not need the questions or they may need to be rephrased. The administration plan will help you figure out in advance how best to get the survey conducted and the completed survey retrieved from as many of the targeted respondents as possible.

Improving the Effectiveness of Interviews (adapted from Patton 1987)

An interview is not a conversation; rather, it is a series of questions designed to invoke answers in the respondent's own words. An interviewer should not interrupt the respondent (other than to regain control or move the interview along), and an interviewer should not share opinions about the questions or comment on the responses.

The following ten tips can make evaluation interviews more useful.

1. Select the type of interview (e.g., structured, where all questions are asked in the same way and the same order; semi-structured, where order and exact wording are more flexible; unstructured, where only topics are identified) or combination of types that is most appropriate to the purposes of the evaluation. Develop an analysis plan, sampling strategy, and reporting template.

2. Communicate clearly what information is desired and why that information is important. Let the respondent know how the interview is progressing.

3. Remember to ask single questions and to use clear and appropriate language. Check (or summarize) occasionally to be sure you are hearing and recording the respondent's responses accurately. Avoid leading questions.

4. Don't conduct the interview like an interrogation, demanding a response. Ask questions and obtain answers, being sure to listen attentively and respond appropriately to let the person know she or he is being heard.

5. Recognize when the respondent is not clearly answering the question and press for a full response.

6. Maintain neutrality toward the specific content of response. You are there to collect information, not to make judgments about the respondent.

7. Observe while interviewing. Be aware of and sensitive to how the person is affected by and responds to different questions.

8. Maintain control of the interview.

9. Treat the person being interviewed with respect. Keep in mind that it is a privilege and responsibility to learn about another person's experience.

10. Practice interviewing. Develop your skills.

Using Observations Effectively for Evaluation Data Collection

The purpose of conducting observations is to describe a program thoroughly and carefully and in sufficient detail so that users of the observation report will know what has occurred and how it has occurred. Observations involve **looking and listening**. A particular strength of observations is that data are collected in the field, where the action is, as it happens.

Effective observations require careful preparation. It is useful to know what is supposed to happen and develop a protocol that includes cues to remind the observer what to focus on, that facilitates recording what is seen and heard, and that incorporates use of codes to facilitate the analysis process where possible (see Appendix of the *Participatory Evaluation Guide* for examples of observation protocols). Additionally, analysis plans and reporting templates should be developed with any protocols.

During the observation, be sure to use the protocol. At the conclusion of an observation, **always** ask, "*Was this a typical session*?"

Making the Most of Record Reviews Using Data that Are Already Available

Review of program records involves accessing existing internal information or information that was collected for other purposes. Data are obtained from a program's own records (e.g., intake forms, program attendance); from records used by other agencies (e.g., report cards, drug screening results, or hospital birth data); or by adding questions to standard record-keeping strategies (e.g., a question for parents about program value added to an enrollment form).

• Remember, record reviews require that you extract data that were collected for other purposes. Make sure that the use of other records makes sense given your evaluation questions, and that any data

collection limitations and inconsistencies in the existing data will not also cause major limitations to your study.

- Be sure to address confidentiality and access issues long before you actually need the data. How does your source collect and maintain the data? Often, record review data can be retrieved via electronic submission if you know how to ask. Develop standard confidentiality agreements and request protocols as needed. *Beware! Securing access, extracting, and modifying data for use in your study may take considerable time, and the steps are dependent on the progress of others.*

RECORD REVIEW CAUTION

It is rare that you will be able to influence how data are collected by the original users. For example: If you were evaluating a program providing access and support services for high school graduates about to be first-generation college enrollees, and you wanted to follow up their matriculation and progress as undergraduates, you would need access to records from all the colleges your participants ultimately attended. Each would have its own system of tracking enrollment, etc. Or, you might access a central repository of information on college enrollment and retention such as the National Student Clearinghouse. But you must keep in mind that system was developed originally to provide lending organizations with enrollment verifications and deferments of financial aid. Its accuracy is limited by timing requirements for high school graduation and credit accumulation updates, by FERPA blocks and by nonparticipation in the system (more than 92% of post-secondary institutes participate, but others do not).

Data Analysis Basics

Staff members at nonprofit organizations often believe this is where their expertise regarding evaluation ends. It can be helpful to bring in special assistance to conduct data entry and analysis, but it is not required.

The Rochester Effectiveness (REP) and Hartford Building Evaluation Capacity (BEC) projects and others like them have shown conclusively that staff within nonprofit organizations have the ability to do good data analysis, i.e., to become "analysts," using readily available tools (such as Microsoft Excel).

Use the following steps to guide analyses of any data:

1. **Organize and enter data into databases**. Although the term database often is used to mean an electronic file (such as an Excel spreadsheet or an Access, SPSS or customized information management software file), it can also be a neat stack of completed surveys, interviews or enrollment forms, or a handwritten grid that captures results. The important thing is that the data are arranged in such a way that it will be easy to inspect and look for trends.

2. **Clean and verify information to store in electronic databases**. Visually inspect the database and double-check a sample to make sure there were no systematic data entry errors (e.g., review about 10% of the electronic data records to be sure they are the same as the paper records). If needed, calculate minimum, maximum, and mean values to see if there are outliers (i.e., extremes or exceptional cases) in the data set. Determine whether outliers should be removed or included with the rest of the data.

3. **Develop a plan *before* analyzing data**.

- Specify what you need to know in advance. It is helpful to go back to the evaluation questions and reclarify what you need to know and how the results of the analysis will answer those questions.

- Specify *how good is good enough* (i.e., clarify targets and how it will be decided whether results—differences between targets and actual outcomes—are favorable/positive, unfavorable/negative, or neutral).

- Specify how data will be encoded (e.g., when categories like *Excellent* and *Very Good = Positive* will be combined).

- Specify what analytical strategies or calculations will be performed (e.g., frequencies—counts or percentages of events or responses, cross tabulations where two or more variables are reviewed together such as gender and age, or status by year). Clarify how missing data will be handled.

- Clarify how data will be disaggregated. Most of the time, participants in a program are not all the same, and there is a need to look at the outcomes or responses of groups of participants separately or comparatively (i.e., to cross tabulate or look at more than one variable at a time). It is common, for example, to present separate and combined results for those of different genders or age groups or from groups with differing levels of experience with the program.

- Determine whether and which tests and analytic strategies (e.g., chi square tests, t-tests, analysis of variance [ANOVA]) are needed to determine statistically significant change, and whether effect size must be calculated to determine magnitude of change. (Refer to the bibliography in *Evaluation Essentials* for examples of references that clarify the use of different statistical tests. **Keep in mind that most basic**

program evaluation and use of evaluative thinking to inform organizational decisions depends on relevance and meaningfulness rather than statistical tests.

Example 1: An organization served a large group of youth in a summer learning loss prevention program. Pre-program and post-program reading test scores were not significantly different (although on average, post-program scores increased slightly). The program still concluded it had achieved something important because the expectation was that there might be significant decline over the summer.

Example 2: An organization serving female offenders with high rates of recidivism compared the number of days participants remained unincarcerated to the number of days unincarcerated for those in a similarly-sized population with similar profiles that had not received any post-incarceration support. The difference in average number of days participants stayed free was very small, but significant (121 days on average for the target group as compared to 119.5 days on average for the comparison group). However, the organization did not consider this as strong evidence their program worked, because they were trying to promote freedom for more sustained periods, six months or more, in an effort to break the cycle of recidivism. While their participants did significantly better than those with no services, they missed the 180-day minimum target by a substantial amount. In response, the organization went back to their participants, staff, and other consultants to determine what else their program needed to help support participants.

- Determine how to present results, (e.g., as numbers, percentages, in tables, or as graphs). See also the final section of this manual for additional suggestions regarding communication of results.

4. **Develop "dummy" tables and proposed graphics, lists, and outlines for analyzed data**. Share plans for presenting summarized data with others in advance (e.g., table shells with no data in them but clear labels for what will be in the rows and columns).

5. **Identify the most important findings from the data**. Once you have completed the analyses and identified key findings, summarize them, and then use the specific results (e.g., a table or list of data) to clarify the summarized findings. The most important findings are those that answer the evaluation questions. Like evaluation questions, the number of key findings should be limited and focused (i.e., usually between about three and 10 findings are key). Write them as concise sentences.

6. **Present results of the analysis in an orderly, meaningful, straightforward way**. Usually this is done using tables, bulleted lists, or brief narrative summaries. It is the analyst's job to help the targeted audience know what they should pay attention to. Formatting and clear language all help to accomplish that goal.

Things to Avoid When Reporting the Results of Analyses

1. **Including response rates and problems with methodology as part of the findings**. Neither of these are actually *findings*. They belong in a description of data collection methods.

2. **Reporting both numbers and percents**. Most of the time, numeric data are reported as percents because that is what most readers are accustomed to. But be careful not to use percents when there is a small set of data (e.g., if you only interviewed six informants, don't report the percent who responded in certain ways). For data sets of less than about 50 cases, use numbers or meaningful categories (such as *all*, *some*, or *most*).

3. **Listing in a sentence or a table all of the response choices for every question on a survey, record**

review protocol, or interview guide. It is the job of those who analyze the data (be they internal staff or a consultant brought in specifically for this purpose) to bring the readers' attention to what is important.

For example, rather than stating that 15% of the participants made no gains during the program, 20% of participants made 1–5 gains, 35% of the participants made 6–10 gains, and 30% made more than 10 gains, point out the critical facts. Almost two-thirds of the participants (65%) made more than 5 gains; OR a total of 85% made at least some gains; OR a total of 15% of participants made no gains during the program, and only 30% made more than 10 gains. Use a predeveloped analysis plan to help guide summarization of findings. Present data in such a way that readers can calculate findings on their own.

4. **Reporting results with excessive precision**. Most of the time it is acceptable to round to the nearest whole number when reporting percentages. For example, 68.473% of the participants would become 69% of the participants. Sometimes, such as when there are data from multiple sites that are similar, or when looking at cumulative change that varies incrementally, additional precision can be helpful to distinguish differences (e.g., first test score = 25.2; second test score = 27.5; third test score = 29.3; fourth test score = 33.5; final test score = 35.2).

5. **Keeping results in the same order as they appeared on survey or interview protocols**. It is the job of the analyst to order things in the best way to clarify the findings—you are **not required** to present things in the order you asked them.

6. **Compartmentalizing results**. It is distracting for a reader to see *what the surveys told us, what respondents said in interviews, how our participants did on tests*, etc. Combine the information from various data collection sources to clarify relevant findings about the subject of your evaluation.

It is the analyst's job to show how information from a variety of data collection sources revealed important, actionable findings, or to clarify when and why there may be discrepancies in data collected via different sources.

For example, the evaluation may ask participants and some significant others about the impact of a program on them, e.g., teenagers in a program and their parents or school teachers. The findings could show that teenagers thought the program was critically important while their parents weren't aware of this at all, or the reverse, or you could find that parents corroborated or expanded on the importance of the program described by the participants. Both are important and actionable findings.

7. **Feeling compelled to use all of the information collected**. Sometimes evaluation data collection does not work out as planned. In fact, research has shown that most evaluation studies vary from original plans. Don't use partial information or results of data collection activities that were problematic.

For example, if you administered a batch of surveys and a few were filled out incorrectly or by people who should not have been in the target population, eliminate them. Explain, if needed, why the information was discarded, but concentrate on findings that answer the key questions of the evaluation work.

8. **Including any action steps or conclusions that are not clearly developed from the findings**. Any conclusions or summary statements should be illustrated by actual data collection results.

9. **Introducing totally new topics into the report in the final sections**. Do not use the report to explain why the design was changed, what wasn't done, or what should be happening with a program, regardless of the findings presented in the report.

For more information about how to strengthen evaluation capacity in your organization, consider reviewing available texts on the subject. The Bruner Foundation's *Participatory Evaluation Essentials: an Updated Guide for Nonprofit Organizations and Their Evaluation Partners* (Baker and Bruner 2010) contains specific guidance for conducting evaluation and building evaluation capacity as well as a bibliography of other evaluation resources. (Appendix 8 in this guidebook also includes details on commissioning and paying for evaluation.)

Organizations that use evaluative thinking:

- Regularly conduct evaluations that include attention to characteristics, activities, and outcomes of selected programs.

- Involve multiple evaluation stakeholders including staff, organization leaders (sometimes even board members), and clients (as appropriate) in developing and revising evaluation plans as well as collecting and analyzing evaluation data.

- Share results of evaluations, including findings about client outcomes (as appropriate), with multiple evaluation stakeholders including leaders, staff, clients, board members, and funders.

- Use results of evaluation to generate knowledge, render judgments, inform decisions, and improve programs where necessary.

- Insure that key staff with evaluation expertise address the organization's evaluation needs and that there are staff members whose jobs or components of their jobs are dedicated to evaluation.

- Hire evaluation consultants when needed (see Appendix 8 for details).

- Provide or obtain training in evaluation for program staff members and make sure that the training is current, well-delivered, and provided for enough staff members to insure that evaluation use is a standard practice.

BUILDING EVALUATION CAPACITY PROJECTS

In projects devoted to evaluation capacity building (e.g., Rochester Effectiveness Project [REP] conducted by the Bruner Foundation 1996–2003; Building Evaluation Capacity [BEC] conducted by the Hartford Foundation for Public Giving, 2006–present; Evaluation Institute [MWEI] conducted by the MetroWest Health Foundation in Framingham, Massachusetts), all of the information in the *Participatory Evaluation Essentials* guidebook was delivered in a series of comprehensive multi-hour training sessions. Each session included a presentation of information, hands-on activities about the session topic, opportunities for discussion and questions, and homework for trainees. By the end of the training sessions, trainees had developed their own evaluation designs. For REP and BEC, the trainee organizations also implemented their designs during an additional 10 months of evaluation coaching and review. As projects concluded, trainees summarized and presented the findings from their evaluations. The REP nonprofit partners and other nonprofit provider trainee organizations in Hartford agreed that the up-front training helped prepare them to do solid evaluation work and provided opportunities for them to increase participation in evaluation within their organizations. We recommend this approach for those who are interested in building evaluation capacity.

Representatives from the trainee organizations in Hartford and Framingham also reported that their increased evaluation capacity and opportunities to assess evaluative thinking in their organizations helped them understand and make application of evaluative thinking in multiple areas of their organizational work. For example, trainee organizations shifted their organizational structure to insure that there are key staff with evaluation expertise to address the organizations' needs and that there are staff members whose jobs or components of their jobs are dedicated to evaluation.

The Bruner Foundation is developing an automated matrix/guide that describes key features such as context, convener, timeframe, deliverable, and the cost of Evaluation Capacity Building projects it has been involved with over the past 15+ years. The guide is designed to help both grantmakers and nonprofit organizations consider different strategies for enhancing evaluative capacity.

5. Board Members, Evaluation, and Evaluative Thinking

As those responsible for the governance of the organization, board members have a significant role to play in ensuring that the organizational leaders use data-driven decision making. Therefore, they can have a significant role to play in supporting, encouraging, and sometimes driving evaluative thinking in order to strengthen the organization. Nonprofit leaders and staff are routinely involved with multiple stakeholders. It is our belief that evaluators, funders, program providers and their board members can and often should be meaningfully engaged in evaluation.

Involvement of Board Members in Evaluation

Board members, though often seen only as recipients of evaluation information, can also play important roles in the process of evaluation. Involvement of board members who collaborate with staff, organization leaders, and/or evaluators can help to:

- Contribute to a sense of shared responsibility for and ownership of evaluation

- Increase the quality of evaluation by including different perspectives

- Provide opportunities to discuss and clarify expectations for the evaluation

- Provide opportunities for board members and staff to interact

- Reduce fears and misunderstandings about the evaluation

- Surface or highlight multiple perspectives that exist (e.g., among staff, between staff and program managers or executive leaders, between direct clients and their caregivers)

This requires a collaborative approach by selected board members and usually additional training for them regarding effective evaluation practice. To facilitate meaningful involvement of board members:

- Determine who from the board should be part of the stakeholder group

- Make sure selected board members understand effective evaluation practices (don't assume they all will understand, as it is still a relatively new field; don't assume that management or even research expertise is equivalent to knowledge about evaluation)

- Clarify specific roles for individual board members based on the evaluation plans, board member capabilities, and timeframe

- Be clear about and get agreement about time commitment, responsibilities, and working in partnership with leaders, staff, and evaluators (avoid creating scenarios where board members' interests, ideas, or responses are deemed more valuable than those of others)

- Specify ways and a timeframe for reporting progress to the whole board

When and How Board Members Can Be Involved in Evaluation

Depending on their skill sets and availability, board members can be part of a group that helps focus and conceptualize the evaluation, interpret data and develop action steps, and communicate about evaluation and evaluation results.

During the evaluation planning stage, board members can help:

• Clarify the evaluation questions and purpose for the evaluation

• Help select evaluation consultants

• Review and refine program logic model or theory of change including clarification of outcomes identification, specification of indicators, and setting of targets

• Select data collection strategies, specify timelines, and determine products

• Review data collection instruments, administration and analysis plans, and report outlines

While the evaluation is being conducted, board members can help:

• Hear and reflect on updates on evaluation progress and preliminary findings

• Begin to think about communications plans

• Connect evaluation team members to other resources or expertise as needed

When the evaluation is completed, board members can help:

• Discuss and interpret results from summarized findings and develop action steps

• Determine audiences and formats for evaluation reporting

Remember again, all members of the board do not need to have comprehensive knowledge about evaluation, but it is helpful for everyone to learn or know basic information, and for those involved more seriously in evaluation actions to learn or know specifics like those identified in Sections III and IV. The following are requisites. **Board members must understand and have consensus on**:

• The working definition for evaluation and evaluative thinking, and what is expected of organization representatives engaged in participatory evaluation

• The purpose of the specific evaluation they are working on

• Reasonable outcomes and related indicators and analytical strategies

• How to commission and pay for evaluation

• How to interpret and report evaluation findings

• How evaluation findings will be used

• Why evaluative thinking is important and how they can incorporate evaluative thinking into their own work

Because it is important that stakeholders are "on the same page," the following recaps important information summarized in Section III.

- All evaluations are **partly social**, because they involve human beings; **partly political**, because knowledge is power and decisions about what gets asked and how information is used are political decisions at the organizational level; and **partly technical**, because choices always have to be made about how to collect and analyze data (Herman, Morris, Fitz-Gibbons 1996).

- Having a clear evaluation design ensures better, more timely, information.

- Both **qualitative** (descriptive) and **quantitative** (numerical) data can be collected using multiple methods (e.g., observations, interviews, surveys, statistical analyses of practical assessments from review of program records). Although there has been much debate about which strategies and types of data are best, current thinking indicates that **both qualitative and quantitative data are valuable and both can be collected and analyzed rigorously**.

- There is no "silver bullet" or "one size fits all" to address evaluation needs. Different evaluation needs and available resources dictate different designs, types of data, and data collection strategies. Comparison groups, experimental designs and even "pre and post" testing are not required or even appropriate in many program evaluation designs. (Be sure to review Appendix 5 on the differences between evaluation and research for more clarification about evaluation standards.)

Evaluations are typically conducted to accomplish one, two, or all of the following:

- To account for or document grant or program actions or expenses

- To generate new knowledge

- To render judgments

- To inform data-driven decision making and facilitate improvements as needed

It is critical to carefully specify what program (or program component) is to be evaluated, why the evaluation is being conducted, and what will be done with the findings (see also Section VI for more information about evaluation findings).

Review all sections in this guidebook (especially III and IV) together with board members to help prepare them for their roles.

Working With Your Board to Define Program Outcomes, Indicators, and Targets

Clarifying outcomes, indicators, and targets for programs is some of the hardest work of evaluation and a common focus of stakeholder, including board member, involvement (see Section III). **Keep in mind the following**:

- Outcomes, especially long-term outcomes, should not go beyond the program's purpose (i.e., don't project educational outcomes for an employment and training program).

- Outcomes should not go beyond the scope of the target audience (i.e., don't project change throughout the county if you are only serving a small proportion of county residents in a particular neighborhood).

- Avoid holding a program accountable for outcomes that are tracked and influenced largely by another system, unless there is meaningful interaction with that system regarding outcome change (e.g., don't hold an afterschool program accountable for the outcomes of students at school, unless the afterschool and day school programs are integrated).

- Do not assume that all subpopulations will have similar outcomes (e.g., outcomes may be very different for those with longer program experience; different subgroups may require different targets).

- Be sure to measure outcomes on a timetable that corresponds to logical projections of when outcomes could be accomplished. (For example, do not assume you will be able to report about school outcomes until after the end of the school year.)

- Most outcomes have more than one indicator, and indicators may not capture all aspects of an outcome. Identify the set that is believed (or can be agreed) to adequately and accurately signal achievement of an outcome. For example:

Outcomes	Indicators
Initial: Teens are knowledgeable about prenatal nutrition and health guidelines.	Program participants are able to identify food items that are good sources of major dietary requirements.
Intermediate: Teens follow proper nutrition and health guidelines.	Participants are within proper ranges for prenatal weight gain. Participants abstain from smoking. Participants take prenatal vitamins.
Longer Term: Teens deliver healthy babies.	Newborns weigh at least 5.5 pounds and score 7 or above on the APGAR scale.
Participants will be actively involved in afterschool activities.	At least 500 students will participate each month. Students will attend 70% or more of all available sessions. At least half of the participants will participate in 100 or more hours per semester.
Clients show slowed or prevented disease progression at six and 12 months.	A total of 66% or more of all clients will have: — Sustained CD4 counts within 50 cells — Viral loads <5000

- To make judgments about a program or facilitate improvement of a program, you need not measure all indicators for all participants. Selection of sample data for key indicators can provide ample information for decision making (see also *Participatory Evaluation Essentials: an Updated Guide for Nonprofit Organizations and Their Evaluation Partners, Baker and Bruner 2010*, for more information about sampling strategies and sample size).

- Set targets in advance based on: best professional hunches, external standards (when they are available), and past performance (when baseline or initial data are available). Do not agree to targets that are unrealistically high or embarrassingly low.

Board Involvement in Securing Evaluation Consultants

Evaluation assistance can be obtained from independent technical assistance or evaluation consultants, evaluation or other technical assistance consulting firms, and sometimes universities with graduate programs that include training or projects in evaluation (especially program evaluation). Along with key organizational decision makers, board members can be collaboratively involved in the identification and even selection of evaluation consultants.

Before hiring any consultant or organization be sure to find out whether they have:

- Experience with evaluation, especially with non-profit organizations

- Basic knowledge of the substantive area being evaluated

- Good references (from sources you trust)

- A personal style that fits with your organization's working style

For more details about commissioning evaluation, see Appendix 8.

What About Governance and Evaluative Thinking?

Governance is the primary role for the board of directors in nonprofit organizations. In addition to the roles in evaluation described above, board members can and should incorporate evaluative thinking into their governance roles.

As stated in Section I, *in organizations where there is evaluative thinking*:

- The board is aware of and understands evaluations of the organization's programs, initiatives, and strategies.
- The board uses appropriate data in defining the goals, work plan, and structure to develop plans summarizing strategic direction.

- The board regularly evaluates progress relative to own goals, work plan, and structure.
- The relationship between organization mission and plans for strategic direction are assessed.
- There is a systematic process and timeline for identifying, recruiting, and electing new board members.
- Specific expertise needs are identified and used to guide board member recruitment.
- The board regularly (e.g., annually) evaluates the executive director's performance based on established goals and work plan.
- The board assesses the organization's progress relative to long-term financial plans.
- The board includes appropriate evaluation results in the assessment of the organization's progress.

6. Using Findings

Effective evaluative thinking and evaluation depend on successful, timely communication of results and ultimately use of findings to inform decision making. The following are important considerations related to information sharing to promote findings use.

Uses and Packaging for Evaluation Findings

Most evaluation findings are presented in formal evaluation reports. These are usually written documents that include a cover page; table of contents (optional); and sections describing what was done (methodology), what was found (results), what it means (discussion/interpretations) and what can be done (action steps/recommendations). These can be produced with word processing or presentation software (such as PowerPoint). Internal and external evaluation reports can be developed, as well as a brief executive summary with minimum descriptions of the work and a discussion and presentation of key findings. It is common for multiple reports to be developed when there are multiple stakeholder audiences such as staff, board, clients, or a larger community.

In addition to written reports, evaluation findings and results of evaluative inquiry are also commonly disseminated via oral presentations, again differently to different audiences, and sometimes via more creative means such as display posters, flyers, and brochures. Any of these dissemination forms can also be posted to websites, Facebook pages, or blogs and linked electronically to other informative sites or reports. Organizations that regularly use evaluative thinking will do the following regarding development and dissemination of evaluation findings:

- **Develop a reporting strategy in the evaluation design stage**. Determine who needs to see and hear findings—i.e., who the audiences are. (It may also be helpful to establish general guidelines about reporting—e.g., executive summaries are always developed, or never developed, and findings always go to involved program staff before they are presented either in writing or orally to other stakeholders).

- **If draft communication is needed, be sure to determine who should be involved**, how they should respond to draft information, and how much time is available for response. You also need to figure out what you will do with the feedback (e.g., incorporate it, make changes based on it, or just note it).

- **Don't assume reporting is only about compliance** (i.e., reporting out to stakeholders such as funders, board members, or senior managers). For reporting to exemplify evaluative thinking, it must inspire action.

- **Figure out the best way to communicate evaluation findings to various audiences**. For example, if written reports are never used and unnecessary for documentation, find other ways—such as oral presentations—to present findings. Or, if board members always want both a written report and an oral presentation, be sure to factor in timing and

skills required to get both of those accomplished. Don't forget about translation if any versions need to be shared with audiences who use languages other than English.

- **Plan for and allocate both time and financial resources to effectively communicate findings**. Development of reports and/or presentations takes both time and skill. It is often the most time-consuming aspect of evaluation. It should **not**, however, be seen as a task that should always be farmed out. The process of developing written or oral reports is usually very instructive regarding what the findings are and what they mean.

- **Revisit the planned reporting strategy and make necessary changes**. Once evaluation projects are underway, it is common to recognize new audiences and/or additional reports that should be developed.

As findings are readied for reporting, also consider the following:

- Do the findings accurately reflect the data that were collected and analyzed?

- How might the interpretation be inaccurate?

- Are there any unintended consequences that might result from sharing the findings?

- Are there any missing or overlooked voices?

Lastly, timing and venue are also important considerations. Along with audience identification and presentation strategy selection, those responsible for communicating evaluation or inquiry results should consider whether there are:

- Natural opportunities for sharing findings such as regularly occurring or preplanned meetings focused on the evaluation or inquiry subject

- Opportunities to host a special convening for the purpose of presenting findings

- Opportunities during regular work interactions (e.g., clinical supervision, staff meetings, board meetings) where selected findings could be presented

- Informal discussions including review of results

"Ate" Steps to Promote Effective Information Sharing

1. **Deliberate**—Spend time with people close to the evaluation work and confirm the findings. You must convince yourself/ves first.

2. **Anticipate**—Determine how you want to use the findings and what value might be derived from the findings.

3. **Investigate**—Once you have findings, test them with key stakeholders. They will shed light on perceived value of the findings.

4. **Calibrate**—Develop a result-sharing mechanism that can convey the message you want to convey to your chosen audience.

5. **Illuminate**—Remove any unnecessary details and highlight the key findings.

6. **Substantiate**—Take a step away from the work and come back to it later with fresh eyes. Ask yourself, "Do the findings still resonate?"

7. **Annotate**—Proofread the final draft. Misteaks can distract from results.

8. **Communicate**—Share the results!

Modified with permission from a presentation made by Howard Walters, OMG Center for Collaborative Learning, to nonprofit organization evaluation capacity building trainees

Evaluation Reports Revisited

When a full evaluation is conducted (e.g., of a specific program), the evaluation results are usually summarized into an evaluation report. The following provides guidelines for completing a formal evaluation report:

1. Determine the needs, purposes, and probable audiences for evaluation reporting. (Remember key stakeholders. How can results best be reported to clients, staff, funders, and others?)

2. Develop a report outline (be sure to include the components identified in the next section).

3. Determine the desired reporting formats (written document, electronic document, written or electronic presentation materials, executive summaries, consumer reports, etc.) and develop a report production timeline with writing assignments.

4. Assign someone (or a team) the task of primary author(s) and the responsibility of final production. Assign others specific writing assignments as needed.

5. Develop a report production timeline.

6. Develop a dissemination plan. (It is common and desirable to develop multiple evaluation products for different stakeholders. For example, staff may need a very detailed report while executive leaders or boards need only a summary of key findings.)

7. Share the report outline, audience list, suggested reporting formats, proposed timeline, and dissemination plan with key stakeholders.

8. Revise the report outline and all other report plans to incorporate key stakeholder suggestions.

9. Complete writing assignments so that all the report sections (according to the revised report outline) can be compiled by the primary author(s).

10. Combine all sections and develop the first report draft. Share the draft with appropriate stakeholders.

11. Make revisions as needed and produce all necessary versions of the report.

Components of a Strong Evaluation Report

A strong evaluation report should include the following:

• Program description

• Purpose of the evaluation

• Evaluation questions

• Description of actual data collection methods used

• Summary of key findings (including tables, graphs, vignettes, quotes, etc.) and a discussion or explanation of the meaning and importance of key findings

• Suggested action steps based on the findings

• Next Steps (for the program and the evaluation)

• Issues for Further Consideration (loose ends)

When **writing an evaluation report**, be sure to follow the guidelines in the previous section or at least develop an outline first and pass it by some stakeholders. When **commissioning an evaluation report**, ask to see a report outline in advance. When **reviewing others' evaluation reports**, don't assume they are valuable just because they are in a final form. In all cases, review the report carefully for the above, and determine whether the report achieves its purpose.

Important Things to Remember About Report Writing

• Follow the report outline described above. Feel free to be somewhat flexible with the order, but don't leave out whole sections. Make sure findings are featured prominently.

• As suggested in the guidelines, make an internal outline including who is responsible for which sections. Be sure to leave time for stakeholders to help you with editing and making revisions.

• Be economical in decisions about what to include in the report. Shorter is better. Avoid excessive use of jargon.

• Think, in simple terms, about what you are trying to say, and then write that. Use complete sentences and standard English grammar conventions. You can use bulleted statements, but be sure your reader can follow your logic.

• Use tables and graphs to help illustrate findings. All tables and graphs must have titles, labels and legends, or footnotes so that they stand alone.

- Use quotes and vignettes or snippets from field notes to illustrate the findings. Remember, quotes should have quote marks around them and be attributed to the speaker or writer (e.g., "a participant"). If you are presenting field notes, be sure they are clearly identified (e.g., show the date and time the observation or interview were conducted) and in context.

- Use formatting to increase clarity. Headers and sections will help the reader know what is happening in your report. Be consistent about where and how they appear (centered, bold, underlined, side headings, etc.). Number the pages. When generating a draft, think about double-spacing.

- Be consistent in the use of language, capitalization, punctuation, etc. For the most part, evaluation reports should be written in the past tense—only report what you actually did and what you found. The action steps or Issues for Further Consideration sections can include references to future actions.

- Read and re-read your work—if you can't understand it, chances are others won't either.

Preparing Findings to Share

Evaluation findings often include both positive and less positive (negative) information. It is rare when program results are totally good or exceptionally awful. It is useful to think with stakeholders on the front end, before results are available, about what to do with different possible results. The following are steps to employ when results are negative, inaccurate, inconclusive, or positive.

What Should Be Done If Evaluation Results Are Negative?

Explain with as much certainty as possible what the results are and what they mean.

- If targets have been missed, clarify how negative the results are. *For example: If it was expected that 85% of your participants would remain in the program through the six-month period, but only 40% did, it should be explained that rather than having most participants complete the program, most did not.*

- If possible, use available data to explain why the negative results occurred.

- Don't blame bad program results on bad evaluation data collection techniques unless the evaluation was badly designed or implemented. If so, redo it and then report about results.

- Clarify the next course of action based on the results (e.g., the program will be discontinued, staff will change, more staff training will be incorporated, the target population will shift, etc.).

- If possible and useful, clarify what did work and who it worked for—so that can be built upon in future efforts.

- After summarizing findings, as described in the section on basic data analysis, the analyst must try to give the stakeholders some overall judgments. Avoid the milquetoast approach of throwing in a good finding for each negative. Clarify if the findings are mostly positive, or mostly negative, or mixed, and then use examples to support the statement.

What Should Be Done If Evaluation Results Are Inaccurate Or Inconclusive?

Inaccurate—determine whether inaccurate results have been caused by faulty data collection, faulty data analysis, or other evaluation missteps. If a report is released with information that is later deter-

mined to be inaccurate, errata can be released or reports can be recalled. It is uncommon to complete a full report when it is known that results are inaccurate. Strategies for communicating with stakeholders about why results will not be available should be undertaken.

Inconclusive—sometimes results do not indicate with clarity what judgments can be made or what next steps can be undertaken.

• Present results in an unbiased fashion and indicate that no conclusions could be drawn from the information.

• Allow the reader to determine the significance of the information.

• If there is concern that the lack of conclusiveness is due to problems with data collection or analysis, develop a plan to correct any evaluation procedures for future efforts.

• If the lack of conclusiveness stems from program challenges (such as major changes in staffing or service delivery), report that carefully and clarify what could be done to remediate the problem in the future.

What Should Be Done If Evaluation Results Are Positive?

As with negative results, explain with as much certainty as possible what the results are and what they mean.

• If targets have been exceeded, clarify how positive the results are. If possible, use available data to explain why you think the positive results occurred.

• Some organizations believe that evaluation **has** to reveal problems and negative facts about programs, so they automatically distrust positive results. Some programs set up designs that are so biased there is no possibility that anything but positive results will occur. Neither condition is valuable. If, after careful data collection and analysis, it is clear that the results are positive, report them and celebrate the accomplishment.

• As with negative results, it is important to clarify the next course of action based on the results (e.g., the program will be continued without changes, new strategies, staff or target populations will be tried, etc.)

• Resist assuming that positive results suggest that the next iteration of the program can be conducted with many more participants (ramped up) or with new or different groups of participants.

• Design careful follow-ups. If possible and useful, clarify what did work and who it worked for so that can be built upon in future efforts. As with negative reports, it is most instructive to report as carefully as possible when things have gone well so that similar programming choices (approach, staffing, target population selection, etc.) can be tried again.

As stated in the previous section, it is important to incorporate evaluation results into a larger organizational communications plan that specifies what messages are important to circulate, as well as when and how that circulation happens. It is also very important to use results of evaluations to inform next steps in the subject programs and to factor program results into the personnel reviews of key staff and executive leaders responsible for the programs as appropriate.

7. Sustaining Evaluative Capacity and Evaluative Thinking

Organizations that regularly use evaluative thinking encourage the extension and use of evaluative capacity throughout their organizations. Organization leaders support evaluation and have key roles. Additionally, multiple staff members are involved in evaluation capacity-building activities (such as training) and in conducting evaluation.

Extending and Sustaining Evaluation Capacity and Evaluative Thinking

Organizations that have enhanced their evaluative capacity also need to focus on sustaining and extending that capacity. The following activities can help accomplish this:

• Apply the knowledge obtained through evaluation training to multiple evaluation needs in the rganization

• Work to ensure that staff understands how evaluative capacity (and evaluation in particular) is relevant to their work and why data collection and analysis are important

• Engage multiple staff in sequential, brief training sessions that address planning, data collection, data analysis and reporting, and have hands-on learning opportunities

• Involve multiple staff in the processes of designing evaluations or evaluative inquiries, collecting data, analyzing data, and summarizing findings

• Assess and continue to model evaluative thinking in all organizational practice

What does it look like when executive leaders and management staff use evaluative thinking?

In addition to helping to spread evaluative capacity throughout the organization, executive leaders and management staff should also incorporate evaluative thinking, whenever possible, into their daily roles. The following are indicators that it is happening.

Leaders and managers in organizations that regularly use evaluative thinking:

• Support and value evaluation and evaluative thinking. Use evaluation findings and results of findings from evaluative thinking assessments in decision making. (*For example: When a program evaluation shows the work has not hit desired targets, actions are taken to modify or discontinue the program; when development opportunities occur, the organization decides whether to respond to a request or RFP based on a review of mission compatibility, organizational capacity and interest in taking on the new program, and past competitive successes with the RFP sponsor.*)

• Educate staff about the value of evaluation and how to participate effectively in evaluation efforts (*e.g., by identifying training needs among staff, developing or commissioning evaluation training for whole staff groups, and making evaluation-related tasks part of the job descriptions of key staff*).

• Motivate staff to regularly use specific evaluation strategies (*e.g., not only to use surveys but also to*

encourage staff to regularly observe programs using a protocol and to conduct interviews with staff and participants).

- Modify the organizational structure as needed to embrace change in response to evaluation findings (*e.g., discontinue programs or program components, add new or different kinds of staff for a program—see sidebar*).

- Foster the use of technology to support evaluation and evaluative thinking. (*Incorporate analytical software tools such as SPSS, SAS or XLSTAT to facilitate analysis of data. Determine needs for proprietary databases additional to those available through Access or Excel; develop analysis plans for collected data; make sure staff has training to use analytical tools such as pivot tables in Excel.*)

- Use data to set staff goals and evaluate staff performance (*e.g., incorporate rubrics into observa-*

tion protocols for certain staff functions and tasks; predetermine the level of service such as numbers of sessions to conduct and the number of person-hours to provide, then compare actual level of service to projected or desired levels).

- Use data to make staffing decisions (*e.g., to decide which staff works on which projects, which staff members are eligible for promotions or advancements or need more assistance and training opportunities*).

- Include attention to evaluation in all management-level succession planning. Leaders and managers should be expected to value and, where possible, be knowledgeable about evaluation (*e.g., have completed a basic evaluation essentials course or participated on a team responsible for evaluating specific programs*).

Evaluative Capacity at Work: Project Example

Afterschool programs in New York City interested in increasing attendance for middle–school-aged participants reviewed the total number of hours for those youth only for the 2005–06 school and confirmed that middle–school-aged youth had the lowest attendance rates of all participants. In response, they decided to do the following before and during the 2006–07 school year.

- Reorganize staff structure to include older, more-seasoned staff dedicated to middle-school programming

- Provide training for staff in early-adolescent development and effective middle-school programming

- Assign staff to teams with only middle–school-aged youth and conduct regular staff team meetings focused on programming and meeting the needs of middle–school-aged youth

- Introduce early evening programming two days per week

- Utilize age-specific grouping so middle–school-aged youth spent the most time in groups with their peers

This required the programs to shift staff teams, increase the amount of meeting time, and hire differently. It also meant programming times for all groups had to shift some to accommodate the changes for middle–school-aged youth.

At the end of the 2006–07 school year, attendance was reviewed again; overall attendance for middle–school-aged youth had increased by 86% (i.e., the average amount of time middle–school-aged youth spent in the program went from about 79 hours to about 147 hours). On surveys, participants, staff, and parents all agreed that staff and program changes had inspired more attendance. This confirmed that their staffing choices had made a difference and both provided evidence to share with stakeholders and targets to assess attendance for the 2007–08 program year.

Thinking About Evaluation-Related Policies

Organizations use policies to clarify their guiding principles and to identify acceptable and unacceptable procedures. In addition to calling on leaders to help sustain evaluative thinking, organizations can also address it through policies and plans. The development of specific, evaluation-related policies about, for example, leadership succession or staff development, can help to extend or sustain evaluative capacity. The following are questions to inform evaluation-related policy development.

Succession (i.e., the transfer of responsibility or power)

• What must the organization do to ensure that a person filling an executive director or management position will be knowledgeable about and support evaluation and evaluative thinking in the organization?

Staff Development

• What evaluation training is needed and provided?

• How often is evaluation training provided and by whom?

• How does the organization ascertain that trainees have effectively learned about evaluation?

Responsibility for Evaluation

• Who conducts which evaluation tasks (i.e., what roles are fulfilled by executive director, managers, staff, others)? Are specific work descriptions provided and time set aside to address evaluation tasks?

• Is there an "Officer in Charge" to ensure quality for, and oversight of, all evaluation work?

• Are consultants used and who oversees their work?

Compensation

• How are staff who contribute to evaluation compensated for their time?

• Are their levels of compensation based on quality of evaluation work?

General

• *Approach*. Is evaluation including multiple stakeholders always strived for? Are there types of evaluation that will not be undertaken (e.g., randomized controlled experiments)?

• *Subjects*. Are there some areas of organizational work that are off-limits?

• *Schedules*. How much evaluation can or should be done in a selected time frame (e.g., one year)?

Organizational Evaluation Plan Development

Developing an evaluation plan is an important first step to conducting effective evaluation. Organization-wide evaluation plans can further clarify how, when, why, and which evaluations will be conducted. The following are suggested steps to inform organizational evaluation planning.

1. **Inventory current evaluation efforts and/or conduct annual reviews of all evaluation work**. Be sure to clarify the status of all evaluations currently underway including what questions are being addressed, how data are being collected and analyzed, when reporting is expected, and to whom reports are being presented.

2. **Identify current evaluation needs**. Determine whether there are programs that need evaluations initiated, revised, or discontinued.

3. **Specify the organizational evaluation plan— account for all programs**.

— Identify which programs are to be the subjects of new comprehensive evaluation. Have staff, along with stakeholders and/or consultants, develop full evaluation plans for those programs. Plans should include: purpose, evaluation questions, identification of outcomes, indicators, targets, selection of data collection strategies, specification of analysis plans, proposed work schedule, budget (if applicable), and deliverables, reports, and dissemination plans.

— Identify which programs are the subjects of ongoing evaluation. Develop full evaluation plans for those programs.

— Identify any other documentation or reporting required for the term of the organizational evaluation plan.

Getting Started With an Evaluation Inventory

An evaluation inventory is a way to document evaluation activity at an organization. Most organizations have evaluation work underway at all times because they are required to and/or they are committed to using evaluation to inform decision making.

To conduct an inventory, start by answering the following questions:

• What evaluations are currently underway and who, if anyone, is helping to conduct them?

• How are data being collected and on what schedule?

• How are evaluation findings being used?

• Who receives any evaluation reports and in what form?

Evaluation Inventory							
Program Being Evaluated	**Key Evaluation Questions***	**Expected Uses**	**Evaluated By**	**Data Collection Strategies**	**Overall Dates***	**Products/ Dates**	**Reporting To:**
			❑ Internal ❑ External (specify who below)	▢Surveys ▢Interviews ▢Observation ▢Record Review	FROM:_____ (month/year) TO:_____ (month/year)		
			❑ Internal ❑ External (specify who below)	▢Surveys ▢Interviews ▢Observation ▢Record Review	FROM:_____ (month/year) TO:_____ (month/year)		

*Obtain full evaluation design for each selected program including work plans and description of outcomes, indicators, and targets.

The following is an example of a chart that could be used to record the results of an evaluation inventory.

Uses for Evaluation Inventories

The results of an evaluation inventory can be used to accomplish the following:

- Inform stakeholders (including board members, funders, staff) about evaluation work that is already being done

- Identify areas where there is insufficient attention to evaluation

- Help reduce duplication of efforts so that the same data are not collected more than once or managed in multiple databases

- Inform decisions about using evaluation resources

- Assist with projections of needed evaluation resources

- Clarify schedules and highlight peak data collection and analysis times

Next Steps: Increasing and Integrating Evaluative Capacity

Organizations that want to enhance their evaluative capacity should consider the following:

1. Conduct an evaluative thinking assessment to identify where there is substantial or less use of evaluative thinking. [The Evaluative Thinking Assessment tool can be found in the Effectiveness Initiatives section of the Bruner Foundation website: *Featured Resources* on the Evaluative Thinking page.]

2. Use the results of the assessment to develop an action plan.

3. Form a task force or working group to champion development of an evaluative thinking action plan. Develop a schedule and create opportunities for the team to interact regularly with other leadership groups (including representatives from the board).

4. Identify specific planned actions for each capacity area where more evaluative thinking is needed and prioritize the list. Be sure questions like the following can be answered.

   — Do the organization's scores match perceptions of evaluative thinking in the organization?

   — In which areas are there strong indications of evaluative thinking? What can be learned from these areas of strength?

   — In which areas are there limitations that could be addressed? What are the priorities?

   — Is it better to tackle the organizational areas with the least evidence of evaluative thinking or address those areas that would be easier to make progress?

   — What changes would produce the most benefit to the organization?

5. As evaluative thinking is assessed and plans are developed to enhance evaluative thinking use, organizations must also stay vigilant regarding evaluation capacity, making sure that leaders promote the use of evaluation, that staff are charged with conducting evaluation, and that sufficient training exists.

Taken together, this should lead to the integration of evaluation skills and evaluative thinking into multiple organizational practices, ensuring both stronger programs and stronger organizations.

Please visit the Effectiveness Initiatives pages of the Bruner Foundation website, http://www.evaluativethinking.org/, for tools that can help you conduct an evaluative thinking assessment of your organization and move from assessment to action plans.

Appendix 1: Useful Evaluation Terms

Assessment: Synonym for evaluation, but often used to refer to a technique (e.g., practical assessment) or a mini-study.

Benchmarks: Performance data used for comparison purposes. They can be identified from your program's own prior data or relative to performance in the field.

Compliance/Monitoring: Type of evaluation where evaluation questions are focused on adherence to pre-specified procedures.

Comparison Groups: Nonparticipants who are identified as a reference for comparison (e.g., individuals at different sites).

Control Groups: Nonparticipants who are usually identified in the use of an experimental design, ideally on an over-subscribed program (i.e., where there are more participants than slots). The treatment or experimental group actually participates in the program and the control group, although eligible and similar, does not receive or participate in the program. Results of treatment and control group outcomes are then compared to determine program contribution to outcomes.

> *** WARNING ***
> **Comparisons must be conducted very carefully.**

Extrapolation: Modest speculations on the likely applicability of findings to other situations under similar, but not identical, conditions. Extrapolations are logical, thoughtful, and problem-oriented rather than purely empirical, statistical, and probabilistic.

Formative Evaluations: Focus on ways of improving and enhancing programs and are conducted in the early or ongoing stages of a program.

Generalize: To assign qualities based upon group membership, or to make inferences about groups or programs based on the outcomes of a sample or subset of members.

Goals: Conditions (usually broad) that programs are working toward (e.g., to promote well-being).

Indicators: Observable, measurable characteristics of changes that represent elements of an outcome (e.g., normal birth weight is an indicator of a healthy baby outcome).

Needs Assessments: Determine whether existing services are meeting needs, where there are gaps in services, and where there are available resources. These are often conducted prior to initiation of an evaluation or in response to evaluation findings.

Objectives: Something that is worked for or strived for, that can be observed or measured.

Outcomes: Results for participants, during and/or after participation in a program.

Outputs: Products of a program's activity (e.g., # of sessions held, # of participants served).

Qualitative Data: Consist of detailed description of situations, events, people, interactions, and observed behaviors; direct quotations from people about their experiences, attitudes, beliefs and thoughts; and excerpts or entire passages from documents, correspondence, records, and case histories. Qualitative data collection methods permit the evaluator to study selected issues in-depth and detail and typically to produce a wealth of detailed data about a much smaller number of people and cases.

Quantitative Data: Come from questionnaires, tests, standardized observation instruments, and program records. Quantitative data collection methods permit the complexities of the world to be broken into parts and assigned numerical values. To obtain quantitative data it is necessary to be able to categorize the object of interest in ways that permit counting.

Random Assignment: A technique that allows program providers to randomly divide participants into treatment (those who get services) and control groups (those who don't).

Appendix 1: Useful Evaluation Terms

Reliable Measures: Those that can be repeated under similar conditions.

Research: In social science is also a systematic collection of information, but it is undertaken to discover new knowledge, test theories, establish universal truths, and generalize across time and space.

Summative Evaluations: Are aimed at determining the essential effectiveness of a program. They are especially important in making decisions about terminating, maintaining, or extending a program.

Triangulation: Multiple streams of information obtained by either collecting different kinds of data about the same subject; using different workers to complete the same tasks; using multiple methods to obtain data; and using multiple perspectives to analyze information.

Valid Measures: Those that accurately measure what they are intended to measure. (Warning—this is difficult to test. For most social and behavioral variables, no agreed upon testing standards exist).

Appendix 2: Evaluative Thinking Assessment

ORGANIZATION MISSION	Assessment	Priority
a. The mission statement is specific enough to provide a basis for developing goals and objectives	1	
b. The mission is reviewed and revised on a scheduled basis (e.g., annually)with input from key stakeholders as appropriate)	1	
c. The organization regularly assesses compatibility between programs and mission	1	
d. The organization acts on the findings of compatibility assessments (in other words, if a program is not compatible with the mission, it is changed or discontinued)	1	
Comments:		
Please proceed to the next worksheet		

STRATEGIC PLANNING	Assessment	Priority
a. There is a formal process for strategic planning	1	
b. Using evaluative strategies such as interviews and surveys, input is obtained from key stakeholders (staff, board, community, and clients) about strategic direction where appropriate	0	This is a lower priority this year.
c. Activities related to strategic process are assessed at least annually	0	This is a high priority this year
d. Activities related to strategic process involve key stakeholders (staff, board, community, and clients) where appropriate	0	This is a lower priority this year.
e. Strategic plans inform decision making	1	
Comments:		
Please proceed to the next worksheet		

Appendix 2: Evaluative Thinking Assessment

GOVERNANCE		Assessment	Priority
a.	Board uses evaluation data in defining goals/workplan/structure to develop plans summarizing strategic direction	0	This is a high priority this year
b.	Board regularly evaluates progress relative to own goals/workplan/structure	1	
c.	There is a systematic process for identifying, recruiting, and electing new board members	1	
d.	Specific expertise needs are identified and used to guide board member recruitment	0	This is not a priority at all
e.	The board regularly (e.g., annually) evaluates the executive director's performance based on established goals/workplan	1	
f.	The relationship between organization mission and plans for strategic direction are assessed	0	This is a high priority this year
g.	The board assesses the organization's progress relative to long-term financial plans	1	
h.	The board assess the organization's progress relative to program evaluation results	0	This is a high priority this year
Comments:			
Please proceed to the next worksheet			

Appendix 2: Evaluative Thinking Assessment

		Assessment	Priority
	FINANCE		
a.	**The organization has systems in place to provide appropriate financial information to staff and board members**	1	
b.	**The organization monitors its financial information systems to ensure they inform sound financial decisions**	1	
c.	**The organization annually develops a comprehensive operating budget that includes costs for all programs, management and fundraising, and identifies sources of all funding**	1	
d.	**The organization annually reviews the comprehensive budget that includes costs for all programs, management and fundraising, and identifies sources of all funding**	1	
e.	**The organization monitors unit costs of programs and services through the documentation of staff time and direct expenses**	1	
f.	**Financial status of organization is assessed regularly (at least quarterly) by board and executive leaders**	1	
g.	**The organization prepares financial statements on a budget versus actual and/or comparative basis to achieve a better understanding of finances**	0	This is a high priority this year
h.	**The organization periodically forecasts year-end revenues and expenses to inform sound management decisions**	1	
i	**The organization has a review process to monitor whether they are receiving appropriate and accurate financial information whether from a contracted service or internal processing**	1	
j	**Capital needs are reviewed at least annually**	1	
k	**The organization has established a plan identifying actions to take in the event of a reduction or loss in funding**	1	
	Comments:		
	Please proceed to the next worksheet		

Appendix 2: Evaluative Thinking Assessment

LEADERSHIP		
	Assessment	Priority
a. Executive leaders support and value program evaluation and evaluative thinking	1	
b. Executive leaders use evaluation findings in decision making	0	This is a high priority this year
c Plans for executive leadership succession include attention to evaluation–new executive leaders are expected to value and be knowledgeable about evaluation	0	This is a high priority this year
d. Executive leaders educate staff about the value of evaluation and how to participate effectively in evaluation efforts	0	This is a high priority this year
e. Executive leaders motivate staff to regularly use specific evaluation strategies	0	This is a high priority this year
f. Executive leaders modify the organizational structure as needed to embrace change in response to evaluation findings	1	
g. Executive leaders foster use of technology to support evaluation and evaluative thinking	0	This is a high priority this year
h. Management uses data to set staff goals and evaluate staff performance	1	
i. Plans for management succession include attention to evaluation–new managers are expected to value evaluation and where possible are knowledgeable about evaluation	1	
j Staffing decisions (e.g., to decide which staff work on which projects, which staff are eligible for promotions or advancements, or need more assistance) are based on data	0	This is a high priority this year
Comments:		
Please proceed to the next worksheet		

FUND RAISING/FUND DEVELOPMENT		
	Assessment	Priority
a. Organization conducts research on potential fund development opportunities (grants and contracts) and assesses which to pursue	0	This is a high priority this year
b. Organization develops a written fund development plan that clarifies which grants and contracts will be pursued	0	This is a high priority this year
c Organization assesses the written fund development plan to be sure it is being followed and to determine why changes and exceptions are made		
d. Organization revises the plan based on assessments		
e Staff (as appropriate) are involved in writing grant proposals (particularly sections on program design and outcomes)	1	
f The costs and benefits for fund raising events and activities are assessed	0	This is a lower priority
Comments:		
Please proceed to the next worksheet		

Appendix 2: Evaluative Thinking Assessment

EVALUATION		Assessment	Priority
a.	The organization involves program staff, organizational leaders, and clients (as appropriate) in developing/revising program evaluation plans	0	This is a high priority this year
b.	The organization involves program staff, organizational leaders, and clients (as appropriate) in collecting program evaluation data	1	
c	The organization involves program staff, organizational leaders, and clients (as appropriate) in collecting analyzing program evaluation data	0	This is a high priority this year
d.	The organization insures that there are key staff with evaluation expertise to address the organization's evaluation needs	0	This is a lower priority
e	The organization insures that there are staff members whose jobs or components of their jobs are dedicated to evaluation	0	This is a lower priority
f.	The organization provides (or obtains) training in evaluation for program staff members and makes sure that the training is current, well-delivered, and provided for enough staff members to insure that evaluation use is a standard practice	0	This is a high priority this year
g.	The organization hires evaluation consultants when needed	1	
h	Evaluations that include attention to characteristics, activities and program and client outcomes are regularly conducted for organization programs	0	This is a high priority this year
i	Results of program evaluations, including findings about client outcomes, are shared as appropriate with leaders, staff, clients, board members, and funders	1	
j	Results of program evaluation drive continuous improvement of programs	0	This is a high priority this year
Comments:			
	Please proceed to the next worksheet		

CLIENT RELATIONSHIPS			
		Assessment	Priority
a.	Client needs assessments are conducted regularly	1	
b.	Client services are designed in response to determined needs	1	
c.	Client satisfaction is regularly assessed	1	
d.	Client outcomes are regularly assessed	0	This is a high priority this year
e.	Results of client satisfaction assessments are used in developing new programs	1	
f.	Results of client outcome assessments are used in developing new programs	0	This is a high priority this year
Comments:			
	Please proceed to the next worksheet		

Appendix 2: Evaluative Thinking Assessment

PROGRAM DEVELOPMENT		Assessment	Priority
a.	The organization identifies gaps in community services before planning new programs	1	
b.	Findings from program evaluation are incorporated into the program planning process	1	
c.	Multiple stakeholders are involved in developing/revising program plans	1	
d.	Program plans are followed where possible, and there are strategies in place to modify program plans if needed	1	
e.	Each program has a written program plan including a logical formulation	1	
	Comments:	100	
	Please proceed to the next worksheet		

COMMUNICATIONS AND MARKETING		Assessment	Priority
a.	Organization has a marketing and communications plan that is linked to the organization's strategic plan which is used to help the organization achieve its mission.	1	
b.	Multiple stakeholders, including staff and board members and technical assistance providers as needed, are involved in developing the marketing and communications plan.	1	
c	Organization assesses the effectiveness of its marketing and communications planning (i.e., determines whether an accurate message is getting out and whether delivery of the message is furthering the mission of the organization).	1	
d.	Multiple stakeholders including staff and board members and technical assistance providers as needed, are involved in assessing the marketing and communications plan.	1	
	Comments:		
	Please proceed to the next worksheet		

Appendix 2: Evaluative Thinking Assessment

TECHNOLOGY ACQUISITION PLANNING AND TRAINING	Assessment	Priority
a. An assessment process is in place to make decisions about technology maintenance, upgrades, and acquisition	1	
b. Technology systems include software that can be used to manage and analyze evaluation data (e.g., Excel, SPSS)	0	This is a lower priority
c. Technology systems provide data to evaluate client outcomes	1	
d. Technology systems provide data to evaluate organizational management	1	
e. Technology systems are regularly assessed to see if they support evaluation	0	This is a lower priority
f. Staff technology needs are regularly assessed	1	
Comments:		
Please proceed to the next worksheet		

STAFF DEVELOPMENT		
	Assessment	Priority
a. A formal staff development needs assessment is done annually	1	
b. There is a plan for staff development, based on needs assessment data	1	
c. The staff development plan is evaluated	1	
d. There are opportunities for staff to assess staff development training sessions	1	
e Results of staff training assessments influence future staff development	1	
Comments:		
Please proceed to the next worksheet		

Appendix 2: Evaluative Thinking Assessment

HUMAN RESOURCES		Assessment	Priority
a.	Organization has an established personnel performance review process	1	
b.	Performance reviews are used (at least annually) to provide feedback relative to performance expectations	1	
c.	Organization collects <u>and updates</u> information on credentials, training, and cultural competencies of staff	0	This a lower priority
d.	The organization uses results of data collected regarding staff credentials, training, and cultural competencies to recruit culturally-competent staff	0	This a lower priority
e.	The organization uses results of data collected regarding staff credentials, training and cultural competencies to train culturally competent staff	1	
f.	Organization conducts regular (e.g., annual, biannual) staff satisfaction surveys.	1	
g.	Organization uses the results of staff satisfaction surveys to inform modification of policies and procedures	0	This is a higher priority this year
	Comments:		
	Please proceed to the next worksheet		

BUSINESS VENTURE DEVELOPMENT		Assessment	Priority
a.	Organization systematically identifies gaps in community services	1	
b.	Organization assesses whether they have the capacity to bring in new types of business	0	This a lower priority
c.	Organization researches new business venture opportunities	1	
d.	Organization strategies regarding new business ventures are based on capacity findings, results of gap studies, and business venture development research	0	This a lower priority
	Comments:		
	Please proceed to the next worksheet		

Appendix 2: Evaluative Thinking Assessment

	ALLIANCES AND COLLABORATION	Assessment	Priority
a.	Existing partnerships/alliances/collaborations are assessed regularly to determine if they meet organization mission and strategic direction	0	This is not a priority for this year
b.	Planning is conducted to identify additionally needed partnerships/ alliances/collaborations	1	
c.	Existing partnerships/alliances/collaborations are assessed regularly to determine if they are functioning effectively	0	This is not a priority for this year
	Comments:		
	Please proceed to the Summary Table to review assessment scores		

Appendix 3: History of Evaluation

Late 1950s–early 1960s	Evaluation mainly focused on educational assessment, conducted by social science researchers in a small number of universities and organizations.
Mid-1960s, the Johnson Era	The War on Poverty and the Great Society programs of the 1960s spurred a large investment of resources in social and educational programs. Senator Robert Kennedy, concerned that federal money would be misspent and not be used to help disadvantaged children, **delayed passage of the Elementary and Secondary Education Act (ESEA) until an evaluation clause was included**. The resulting bill required submission of an evaluation plan by local education agencies and summary reports by state agencies. As a result, **evaluation requirements became part of every federal grant**. (Early expectations were that *evaluation would illuminate the causes of social problems and the clear and specific means with which to fix such problems*.)
Mid-1970s	Two US-based professional evaluation associations emerged in 1976: the Evaluation Network—mostly university professors and school-based evaluators), and the Evaluation Research Society—primarily government-based and university evaluators. (*In 1985, these two organizations merged to form the American Evaluation Association, AEA, www.eval.org, which now has more than 3700 members worldwide*.)
Throughout 1970s and 1980s	• Growing concerns voiced about the utility of evaluation findings and the use of experimental and quasi-experimental designs. • Huge cuts in social programs during Reagan presidency, resulted in less government involvement, diminished or removed evaluation requirements from federal grants. • Many school districts, universities, private companies, state education departments, the FBI, the FDA, and the General Accounting Office (GAO) developed internal evaluation units.
1990s	Increased emphasis on government program accountability and a movement for organizations to be lean, efficient, global, and more competitive. Evaluation became commonly used not only as part of government mandates but also to improve program effectiveness, enhance organizational learning, and inform allocation decisions in a wide variety of both public and private organizations. **A number of foundations created internal evaluation units, provided support for evaluation activities, or both**.
2000–Present	Increasing and sustained interest in participatory, collaborative, and learning-oriented evaluations. National evaluation associations being established throughout the world.

Preskill and Russ-Eft 2005

Appendix 4: Different Evaluation Purposes Require Different Evaluation Questions

Purpose	Questions
Rendering Judgments [Some need met, some goal attained, some standard achieved. Must specify criteria for judgment in advance.]	To what extent did the program work? To what extent did it attain its goals? Should the program be continued/ended? Was implementation in compliance? Were funds used appropriately, for intended purposes? To what extent were desired client outcomes achieved?
Facilitating Improvements [Using information to monitor program efforts and outcomes regularly over time to provide feedback to improve implementation, to fine-tune strategies, and to make sure that participants are progressing toward desired outcomes.]	What are the programs strengths and weaknesses? How and to what extent are participants progressing toward desired outcomes? Which types of participants are making good progress and which aren't? What kinds of implementation problems have emerged, and how are they addressed? What's happening that wasn't expected? What are staff and participant perceptions of the program? Where can efficiencies be realized? What new ideas are emerging that can be tested?
Generating Knowledge [Conceptual rather than instrumental use of findings.]	How is the program model actually working? What types of interventions are being used? What types of outcomes can be expected? How do you measure them? What are the lessons learned? What policy options are suggested by the findings?

Appendix 5: Distinguishing Between Evaluation and Research

Program evaluation and research are similar, but they have different objectives and data standards. They are also different in size and scope. The following table shows important differences between them.

Program	Evaluation	Research
Objectives	Change and action oriented, aimed at determining impact	Aimed at causality, testing theory
Data	Evidentiary data	Very precise measurements
Numbers of Subjects	Program target groups or samples of target groups (sometimes very small)	Usually study of samples from fairly large populations
Standards	Usefulness, practicality, accuracy, ethicalness	Truth, causality, generalizabilty, theory
Costs	Range from minimal to fairly substantial	Usually high costs
Stakes/Scope	Fairly low stakes, fairly narrow scope (the program)	Very high stakes (e.g., human life or health)
Focus	Whether something is being done well, not necessarily better. Should focus on context, activities, outcomes of participants	Determining the best treatments, solutions, etc. Can include community indicators where appropriate
Use	Should not be conducted unless there is real opportunity to use the results	Sometimes conducted when the use is uncertain

Appendix 6: Making Data Collection Decisions

METHOD	ADVANTAGES	DISADVANTAGES	DECISIONS
SURVEYS (Several commercially available, or unique instruments can be developed)	Easy to quantify and summarize results; quickest and cheapest way to gather new data rigorously; useful for large samples, repeated measures, comparisons between units and norms/targets. **Good for studying attitudes and perceptions—can also collect some behavioral reports**.	Hard to obtain data on behavior, context shaping behavior (attribution). Not suited for subtle, sensitive issues. Surveys are impersonal and difficult to construct. Must address language and administration challenges; must avoid nonresponse, biased or invalid answers, overinterpretation with small samples.	Who gets surveyed (sampling)? How will confidentiality be maintained? Validity of self-assessment? What are standards of desirability? Need for repeated measures - what intervals?
INTERVIEWS (Structured, semi-structured, intercept)	Readily cover many topics and features; can be modified before or during interview; can convey empathy, build trust; rich data; provide understanding of respondents' viewpoints and interpretations. **Good for studying attitudes and perceptions—can also collect some behavioral reports**.	Expensive, sampling problems in large programs; respondent and interviewer bias; non-comparable responses; time consuming to analyze and interpret responses to open-ended questions. Training and protocols required to conduct.	Who gets interviewed (sampling)? How will confidentiality be maintained? Validity of self-assessment? What are standards of desirability? Need for repeated measures—what intervals?
OBSERVATIONS (Participants during program sessions, participants in other settings)	Rich data on hard-to-measure topics (e.g., actual practices, behaviors). Behavioral data independent of self-descriptions, feelings, opinions; data on situational, contextual effects. **Good for studying program implementation and some behavioral changes**.	Constraints on access (timing, distance, objections to intrusion, confidentiality, safety); costly, time-consuming; observer bias, low interobserver reliability; may affect behavior of people observed; hard to analyze, interpret, report data; may seem unscientific. Training and protocols required to conduct.	What subjects will be observed How many at which levels? Need for repeated measures—what intervals?
RECORD REVIEW (E.g., program records, school records, case management records)	Nonreactive; often quantifiable; repeated measures show change; credibility of familiar or standardized measures (e.g., birth weight, arrest incidents, drug test results, staff or parent assessment results); often cheaper and faster than gathering new data; can include data from other independent sources. **Good for determining (behavioral) status**.	**Access**, retrieval, analysis problems can raise costs and time requirements; validity, credibility of sources and measures can be low. Definitions must be determined prior to use, are often externally determined, cannot be customized; need to analyze data in context; limited data on many topics.	Which documents? How can access be obtained? Need for repeated measures—what intervals?

Appendix 6: Making Data Collection Decisions

METHOD	VALIDITY	RELIABILITY	AVAILABLE RESOURCES	CULTURAL APPROPRIATENESS
SURVEYS	**Low** No opportunity for clarification. Participants often choose responses other than those provided. Participants may not want to report private behavior. Participants may not be aware of their own actions, behaviors, or attitudes.	**High** Administration is consistent from one individual to next. Standard response choices provide consistent range of responses. Little opportunity for data collector to influence results.	**Economical** Mass distributed. Costs based on number of mailings, use of phone or mail, incentives.	**Varied** Best for literate, middle–class American-born populations. Particularly bad for immigrants and refugees.
INTERVIEWS	**High** Can clarify questions and probe for more in-depth responses. Personal interaction can establish rapport for open discussion. Focus groups can foster discussion and sharing. Focus groups can clarify individual viewpoints through dialog with others.	**Low** Interviews are unique based on comments of respondents. Different questions and probes likely to be used.	**Moderate** Individual interviews: moderate expense. Focus group: low to moderate expense.	**Strong** Individualized interviews work well when paper formats are threatening or invasive and when behavior or attitudes pose a problem. Focus groups work well when the group opinion is the cultural norm.
OBSERVATIONS	**High** Observers can directly observe behavior that may not be accurately reported otherwise. Observers can directly observe behaviors that have standards developed by professionals or institutions.	**Moderate** Observers need structured protocols for coding their observations. Less structured observer formats reduce reliability because different observers may reach different conclusions.	**Moderate–Expensive** Time is required in order to observe behaviors. This can be mitigated by using "natural observers."	**Moderate** Cultural differences in behavior may be misinterpreted.
RECORD REVIEW	**Low to Moderate** Not really designed to measure, rather to document/record.	**Low to High** Depends on whether there are standards for record keeping.	**Economical** Data are part of the service delivery process and usually already exist. (Use of case records for evaluation requires up front planning). Some issues of access, confidentiality.	**Varied** Depends on service delivery, appropriateness of program. May over- or under-represent certain groups due to bias.

Appendix 7: Evaluation Report Outline

I. **Introduction**

    A. Introduction to agency and/or program (mission, goals, and main activities)

    B. Purpose of evaluation

    C. Evaluation Questions

        1. Clear and concise questions that are in alignment with program mission and goals.
        2. If there are several evaluation questions, group them into categories (e.g., service delivery, staff outcomes, participant outcomes, programmatic impacts).

    D. Report Organizer (optional and un-necessary if report is short)

II. **Methods/Data Collection Strategies**

    A. Description of selected methods (in narrative or table form)—this describes what actually happened during data collection, not what the evaluator set out to or attempted to do.

  *B. Relationship between questions and data collection strategies (usually done as a table)

  *C. Data collection rationale—explanation of why methods were chosen including clarification regarding use of participatory data collection.

  *D. Data collection respondents and missing data—description of the target populations for each data collection activity including why they were selected and whether there is any missing data.

  *E. Data collection challenges

  *F. Description of targets for analysis (this can also be addressed in the findings section).

  * Note Sections b - f, can all be addressed as part of the description of selected methods.

III. **Evaluation Findings**

    A. Summaries of the results of data collection and analysis

    B. Response to evaluation questions (where feasible)

    C. Comparison of findings to targets (where feasible and appropriate)

IV. **Conclusions**

    A. Summary of Key Findings

    B. Final Analysis

      * How does the author understand the data
      * How does the author believe the data will impact the program
      * Strengths and weaknesses of the program as revealed by evaluation findings

    C. Action Steps/Recommendations—what will be done with the report, what could/or should be done with the program

    D. Issues for Further Consideration (any outstanding issues raised by the evaluation)

Appendix 8: Commissioning Evaluation

What to look for in an Evaluator

- **Basic Knowledge of substantive area being evaluated**

- **Experience with program evaluation (especially with non-profit organizations)**

- **Good References (from sources you trust)**

- **Personal style and approach fit (probably most important)**

Before you Commission an Evaluation

- **Talk to a few trusted colleagues** about experiences they have had. If possible also talk to some grant-makers and a professional evaluator. Gather some basic advice and determine if it is relevant for your project.

- **Think about how you will identify evaluators.** Is sole sourcing an option for your project, or will a competitive process be more appropriate or advantageous? There are merits to either approach. If you want to or are required to use a competitive approach, determine how broad the competition should or must be. (Determine answers to these questions.)

  — Will an invitational approach work, or do you want unrestricted competition?

  — Are there any geographic limitations or advantages?

  — Are there tax or business requirements (for-profit./non-profit, private firm, individual, university institute or department, etc.)

  — What sources will you use to inform evaluators about your project and attract bidders (e.g., RFPs posted on your website, in publications, through associations)

- **Determine the best strategy and requirements for proposals.** Can you go ahead with a letter proposal, a letter of interest (LOI)? or an interview only process, or is a Request for Quotes/Qualifications (RFQ), or a full **Request for Proposal** (RFP) best? *Whatever you decide, be sure to make the request specific enough that you make your needs known, but not so detailed that the evaluator or other stakeholders have no real input into how to proceed.*

- **Determine the timeline for finding evaluators.** If you do a full, competitive proposal process, you will need to determine your sequence for announcing your competition, releasing the actual request (RFP or RFQ), conducting a bidders conference or responding to clarification requests, collecting responses, making selections (including possible "best and final" competitions or invited interviews/presentations) and notification. *Remember that good responses, especially those that are written or presented through a meeting take some time to develop—be sure to give potential contractors adequate time.*

- **Determine the format for response.** Do you need a written or oral response, or a combination? What categories of information are required and what additional materials will help. (It is always helpful to provide specific questions of interest—see following—and parameters for response.)

- **Determine who will be involved in making the selection and how.** Do you need external reviewers? How will you manage multiple (conflicting?) reviews? If you host an interview, who needs to participate? What role can/should program staff and leaders play?

Questions to Ask Evaluators (in RFP's, RFQs or interviews)

- **What do you need to know to properly design an evaluation for this program/initiative?** (The evaluator should minimally need to know about the purpose for commissioning the work, as well as details about service delivery or other organizational structure, and scope of the project including timeframe, ballpark program budget and size of

Appendix 8: Commissioning Evaluation

the target population. A smart evaluator will also request other background materials or perhaps even a preliminary visit.)

- **What evaluation questions would guide your effort?** (It may also be valuable to have a preliminary conversation about outcomes and indicators, or specify a whole task where the evaluator works together with program staff to clarify expected outcomes, timeframes, indicators, and important assumptions and contextual issues related to service delivery.)

- **What strategies would you use to address the evaluation questions?** (Be specific about how you would: collect and analyze data involve agency staff, why this approach makes sense or is common, whether there are any standard instruments and why they were chosen.)

- **How will you handle common challenges?** (For example, how will your evaluation design be affected by poor project implementation, which outcome measures would be appropriate if the program is not well implemented? How will you communicate this to stakeholders?) What will you do to insure necessary access to subjects and confidentiality of response?

- **What timeline will the evaluation project operate on?** (Specify in chart or calendar form, when key evaluation tasks will be completed.)

- **Who will conduct the work and what other relevant experiences do they have?** (Identify key staff and clarify their level of involvement—attach resumes and a capacity statement with descriptions of other similar projects. Be sure to get specific directions if web-site reviews are recommended. Ask about supervision if multiple evaluators are involved— who is ultimately responsible for collecting and analyzing data, verifying accuracy and reporting results? For multi-site initiatives, will any local evaluators be involved? How will they and any other staff be trained to conduct specific evaluation activities?)

- **How and when will the findings from the evaluation work be communicated?** What products/deliverables will be developed? (look for multiple products where appropriate) Will the products of the evaluation have any greater usefulness? How are program managers expected to use the information obtained through the evaluation?

- **How will evaluation resources be used to complete this work including professional time, travel, other direct costs, indirect costs?** (Be sure to ask for a task-specific budget.)

Evaluation Resources

- Independent technical assistance or evaluation consultants
- Evaluation consulting firms
- Universities with graduate programs that include training or projects in program evaluation

> Remember, evaluation should not be viewed as in competition with program resources. Evaluations can be funded as a component of a program (using program funds) or as a separate project (using ear-marked or additional funds). A Common rule of thumb is to set aside at least 10% of the cost of the program, for evaluation.

What's the difference between an RFP and an RFQ? Which is best to use?

1) **An RFP is a Request for Proposal**. Responses to RFPs should include specific projections about how a contractor will undertake the work (i.e., data collection strategies, analysis plans, instrument development), what it will cost (including types of projected expenditures), when it will get done, who will work on it, and what products will be delivered. Usually responses to RFPs also include sections where the potential contractor clarifies understanding of the project and the context surrounding the project, identifies key

Appendix 8: Commissioning Evaluation

questions that will guide the evaluation work, clarifies who will be working on the project, what qualifies them and how the project will be managed, and provides specific information about qualifications or previous relevant experiences. Other project-specific sections can also be included depending on needs (e.g., a section on what a contractor thinks are likely evaluation challenges and how those would be handled, a section addressing confidentially and participatory approaches). Any forms or documentation required of contractors are typically appended.

RFPs are more formal and responses are typically lengthier. They are best used under the following conditions: when the grantor is not very familiar with the potential recipients of the request or when there are very specific evaluation needs (especially those dictated by another external source or a project/program model). Grantors should be sure to provide ample time for a full response.

2) **An RFQ can be either a Request for Qualifications or Request for Quote** depending on the nature of the solicitation. Responses to RFQs can include many of the same elements as RFPs, but the information is less detailed, focused on approach and why a specific contractor should be selected to do the work. RFQs typically include sections/questions like the following:

1. What would be your general approach to completing this project?

2. Who from your staff would be involved in this project and why?

3. What are the expected timelines for the work and the expected products?

4. What are the projected costs?

5. What prior experiences qualify you for this work.

Contractors are also typically asked to provide specific references and some are invited to participate in follow-up interviews where they come to provide additional responses to specific questions and to clarify their workplans.

RFQs are somewhat less formal and best used when you have some familiarity with the contractors who might respond, when style fit is paramount, and when you have a briefer time span to make a selection. There is less reading and writing involved with RFQs, but more of an onus on both the grantor and the potential contractors to maximize the interview process to inform the selection decision.

Questions to Consider Before Identifying a Project for Evaluation

1) **Contributions to Mission or Broader Field**

What is the general purpose of the proposed project and how will it contribute to grantmaker's mission? If not, are there other reasons why it should be supported? (This can also be discussed with grantee as long as they are aware of the grantmaker's mission.)

How does this project contribute to the broader field? What are the likely lessons learned?

2) **Implementation and Feasibility**

Does the program target population know about and want to participate in this program? Have all necessary collaborative agreements been secured?

How will you guard against implementation impediments?

3) **Project Design/Staging**

Describe the key components of the project and how they are integrated into the overall project design. Has a reasonable logical formulation been developed for the program?

Appendix 8: Commissioning Evaluation

How will the project be staged over time?

4) **Finances** (ask if necessary)

Please clarify the following details about projected budget:_____

5) **Outcomes & Evaluation**

How is the project expected to impact participants? How will you know when this has happened?

Things To Avoid When Commissioning Evaluation Projects

- Agreeing to fund a program evaluation design that you do not understand.

- Agreeing to fund a program evaluation where disbursement is not attached to deliverables.

- Commissioning a program evaluation on a timetable that is inappropriate for the program.

- Commissioning an overly complicated evaluation design, or one for which there was insufficient stakeholder involvement in the development.

- Assuming that you must ALWAYS measure outcomes.

- Forcing evaluation of outcomes that are inappropriate or beyond the scope of the program.

How Do You Stay Informed About Evaluation Status After You Agree To Support It?

Periodic or Mid-Point Reports

Request that the evaluator make status reports about:

- What evaluation activities have taken place;

- Any data collection or instrument development problems they have encountered;

- Any proposed changes to data collection timelines or strategies;

- Preliminary findings when appropriate.

These status reports should be on a regular schedule that matches with funder needs and the timeline of the evaluation. Compare each status report with proposed activities and timeline. Be prepared to request clarification if there have been challenges (e.g., timeline, access).

Note: do not require evaluator grantees to complete status reports unless there is a real commitment to review them. Be aware that to produce them takes time and costs money.

Project Conclusion

Request a final status report or make it part of the final evaluation report. Compare what was actually done with what was proposed. Assume there will be some differences as not all tasks work as projected.

Appendix 9: Getting Funding for Evaluation

Where does the money come from?

Evaluation, even self-assessment requires financial resources. Usually the cost to do a good evaluation is equivalent to about 10 – 15 percent of the costs to operate the program effectively (i.e., a program with an operational budget of $265,000 per year would require about $25,000 - $40,000 to evaluate for one year). Actual expenditures may be less if the evaluation is more participatory, but other organizational costs may be associated. For example, if program staff are involved in evaluation data collection, there are still resource needs to cover the costs associated with their program work. Most of the funds for evaluation pay for the professional time of those who develop evaluation designs and data collection tools, collect data, analyze data and summarize and present evaluation findings. Other expenses include overhead and direct costs associated with the evaluation (e.g., supplies, computer maintenance, communication, software).

It is bad practice to assume there is a standard fixed evaluation cost (e.g., $10,000) regardless of program size or complexity, and it is dangerous to fund an evaluation project that does not clarify how evaluation funds will be used. The *Participatory Evaluation Essentials* Guide provides instructions for specifying an evaluation timeline, projecting task-specific level of effort (i.e., estimating how much time it will take each person to complete each evaluation-related task) and calculating associated costs. There is also a sample evaluation budget.

There are two ways to project evaluation costs:

1) Identify a reasonable total amount of funds dedicated for evaluation (see cost percentages on the left) and then develop the best evaluation design given those resource requirements (i.e., work backwards from a fixed available amount for evaluation).

2) Develop the best evaluation design for the subject program and then estimate the costs associated with implementing the design. Negotiate design changes if costs exceed available funds.

To obtain funds to use for evaluation purposes:

- **Write evaluation costs into project development budgets** (i.e., add the projected cost of evaluation into the total cost of operating the project). Requests for Proposals (RFP's) often include a section on evaluation. Take it seriously. Rather than inserting the perfunctory fixed amount, try to estimate what it would really cost to do good evaluation* and include that figure as a line-item on the project budget. Then use the money accordingly if the project is funded.

- **Set aside funds for evaluation on a percentage basis into the organizational budget** (e.g., each year set aside 5% of all organizational funds to use for evaluation purposes). *Develop a plan to use these funds.* For example, each year plan to conduct comprehensive evaluation of some programs (e.g., those that are newer or problematic) and then require more limited data collection/review for others.

- **Obtain funds solely for the purposes of evaluation**. Some grantmakers will provide dedicated funds for evaluation, or capacity-building which includes evaluation capacity-building or evaluative thinking enhancement. These are usually accessed through response to an RFP or other proposal development.*

- **Consider sharing and/or pooling resources** with other departments or organizations involved in evaluation. **If you need help to estimate evaluation project costs or to specify a good design, consider contacting an evaluation technical assistance provider. Some evaluation consultants will do this "on-spec" in return for the opportunity to conduct the evaluation if the project gets funded**.

Appendix 10a: Notes on Sampling and Representativeness

Often, surveys are not administered to every participant in a group. Instead some members of the group are selected to respond. This selected group is known as a sample. If the participant group is large, sampling may be advisable, but the following questions must be answered. How will the sample of respondents be identified? How many respondents are needed for valid results? How will the sample be defined and how will representativeness be ensured? The following are some necessary steps .

1) Define the target population to be studied. The term population refers to all possible respondents or subjects of the survey. The population definition must be precisely specified.

2) Decide whether you should try to include all members of the population (census) or to sample.

3) Select a small subset of a population that is representative of the whole population. Unless the population is very small (fewer than 200), sampling is almost always used.

Ways to Select a Sample

There are several ways to select a sample. The most common of these include simple random sampling, stratified samples, convenience samples, and purposeful samples.

- **Simple Random Sampling** approximates drawing a sample out of a hat. The desired number of sample respondents is identified and selected arbitrarily from a randomly arranged list of the total population. Each individual on the list has the same chance for being selected to participate.

- **Stratified Samples** are used when some important characteristics of the population are known prior to data collection. This is also commonly done when participants represent multiple geographic areas, or when there is disproportionate gender representation.

- **Convenience Samples** involve those respondents who can easily be contacted for participation in a survey. While their responses are often enlightening and can be summarized, they should not be generalized to the entire population.

- **Purposeful Samples** include information-rich cases. These can include extreme or deviant cases, maximum variation sampling, typical cases, critical cases (those that can make a point dramatically), and other variations.

Appendix 10b: How Big Should the Sample Be?

The number of program participants will determine whether to include everyone in the evaluation or select a sample, i.e., a smaller group who can represent everyone else and from whom we can **generalize**.

The sample should be as large as a program can afford in terms of time and money. The larger the sample size (compared to the population size), the less error there is in generalizing responses to the whole population—i.e., to all cases or clients a program.

1st RULE OF THUMB: if the population is less than 100, include them all (and strive to get an 80% response rate); if the population is bigger than 100 select a **probability sample**. (See Appendix 10A for sampling strategies.)

Probability samples allow calculation of the likely extent of the deviation of sample characteristics from population characteristics. **Sampling Error** is the term used to refer to the difference between the results obtained from the sample and the results obtained if data had been collected from the entire population.

The objective when drawing samples is to decrease sampling error and to assure confidence that the results are reliable. 2nd RULE OF THUMB: a common standard for program evaluation is 95% confidence with a sampling error of ± 5%. In English that means that you believe that 95 percent of the time the results from your sample, would be off by no more than 5% as compared to the results you would have gotten if you had collected data from everyone.

It is the absolute size of the sample rather than the ratio of sample size to population size that affects the sampling error (Comer and Welch, 1988, p. 192). Sample sizes for varying population sizes and differing sampling error rates have been calculated (see following page). If you wish for more precision use the following calculation (for 95% confidence, 5% error).

$$n=385 \div ((1+ (385/N))$$

Example: If your population is known to have 472 members, then a sample of 212 would be necessary to ensure 95% confidence with no more than 5% error. $385 \div ((1+ (385/472)) = 212$

### Relationship Between Sample Sizes and Sampling Error

- As shown on the next page, when a sample is comparatively large, adding cases provides little additional precision.

- As population sizes increase, the total size of the sample becomes proportionately smaller without affecting error.

- When the population size is small, relatively large proportions are required to produce reasonable error rates.

- A standard proportion (e.g., 33%) will not work as a sampling strategy for varying population sizes.

- 3rd RULE OF THUMB: you must always draw a larger sample than what is planned for because of refusal. To do this, you need to estimate the refusal rate and then factor that into your calculation.

Desired sample size ÷ (1- refusal rate)
= TOTAL SAMPLE.

Appendix 10b: How Big Should the Sample Be?

Sample Sizes (n) @ 95% Confidence, with 3, 5 and 10% Sampling Error			
Population Size (N)	**Sampling Error**		
	±3%	±5%	±10%
100	92	80 (80%)	49
250	203	152 (61%)	70
500	341	217 (43%)	81
750	441	254 (34%)	85
1,000	516	278 (28%)	88
2,500	748	333 (13%)	93
5,000	880	357 (7%)	94
10,000	964	370 (4%)	95
25,000	1,023	378 (2%)	96
50,000	1,045	381 (<1%)	96
100,000	1,056	383 (<1%)	96
1,000,000	1,066	384 (<1%)	96
100,000,000	1,067	384 (<1%)	96

* Adapted from Reisman, 2000, A Field-Guide to Outcomes-Based Program Evaluation
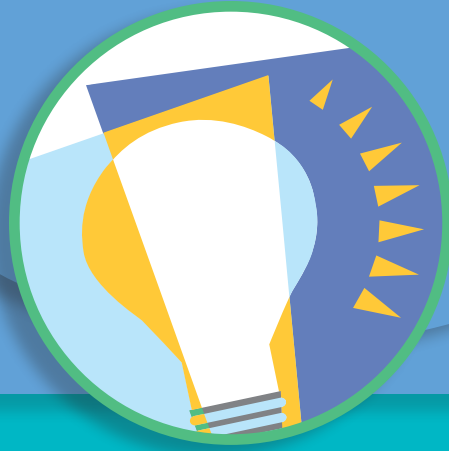
www.brunerfoundation.org

The Bruner Foundation

BRUNER
FOUNDATION

EFFECTIVENESS
INITIATIVES