Chapter 7:
Principles of Evaluation

Chapter Contents

Chapter 7: Principles of Evaluation

Introduction

In previous chapters, we provide information about how to develop evaluation metrics for specific aspects of environmental public health programs. This chapter provides an overview of basic evaluation principles, including:

- Logic models
- Types of evaluations
- Components of evaluation plans

Readers can apply these principles in the planning and implementation of their environmental public health programs to ensure that they are able to document and publicize their successes.

Why evaluate?

Evaluation "involves the systematic collection of information about the activities, characteristics and outcomes of programs, personnel, and products…to reduce uncertainties, improve effectiveness and make decisions with regard to what those programs, personnel, or products are doing and affecting."[65] The benefits of evaluations include the ability to:

- Assess effectiveness and impact
- Determine factors that lead to program success (or failure)
- Identify areas for program improvement
- Justify further funding
- Identify new audiences and applications for projects

When to evaluate?

Evaluations may be undertaken at any time, and they are generally most effective when they are conducted as an integral aspect of the program. Evaluations that are conducted throughout a project's lifespan can provide opportunities for program improvement as the program is evolving rather than after it is complete. Ongoing evaluations also provide an opportunity to adapt the evaluations to address project goals and objectives that may have changed over time. During certain points in a project's lifecycle, there is value in stepping back to examine more fully the operations or impacts of the project. Choosing the right timing depends on the specifics of the project and its particular context. Grantees will likely need to balance many factors, including the evaluation purpose, scale, cost and program resources, when thinking about the timing of an evaluation.

[65] Patton MQ. 1982. Practical Evaluation. Beverly Hills, CA: Sage Publications, Inc. 15.

Metrics in Action 7.1: The **Detroit Community-Academic Urban Research Center (URC)** incorporates evaluations into its overall program planning and development activities. The Detroit URC links a university, eight community-based organizations, a city health department and a healthcare system to identify problems affecting the health of residents of Detroit, Michigan. The partners also promote and conduct interdisciplinary research, which assesses, leverages and enhances the resources and strengths of the communities involved. The URC Board conducts its work in accordance with a set of Community-Based Participatory Research (CBPR) principles adopted by the URC Board that foster, for example, equal participation by all partners in all aspects of the Center's activities and recognition that community-based participatory research is a collaborative process that is mutually beneficial to all partners involved. The 15 member board provides leadership for the group and annually evaluates the partnership and its activities in order to assess the extent to which the partnership is following its key principles of collaboration, participation and equity. The board uses the evaluation findings to build on successes of the program and to share outputs and short-term outcomes with partners. In addition, the findings often lead to changes in board activities, policies or research focus. Conducting annual evaluations allows the Detroit URC to be responsive to short-term changes and to work towards the best possible outcomes.

For more information about the Detroit URC, visit: http://www.detroiturc.org.

Ethical considerations

Because PEPH researchers and evaluators often interact with the community and solicit personal information, it is advisable that they understand their legal and moral obligations to human subjects who participate in research and the evaluation of that research. This understanding can lead to greater trust by their partners and fewer conflicts or misunderstandings down the road. Partners can become familiar with the principles of:[66]

- Ethics
- Confidentiality
- Accountability
- Competency
- Relevancy
- Objectivity
- Independence

For example, university researchers must comply with federal laws and follow the guidelines set out by their institutional review boards (IRBs).[67] When publicizing evaluation findings, partners must remember to keep sensitive information confidential and protect the identities of their subjects.

[66] For more information, see also, U.S. Government Accountability Office (GAO). 2007. Government Auditing Standards, July 2007 Revision. Available: *http://www.gao.gov/new.items/d07731g.pdf* [accessed 16 December 2011]; American Evaluation Association (AEA). 2004. Guiding Principles for Evaluators. Available: *http://www.eval.org/publications/guidingprinciples.asp* [accessed 16 December 2011].

[67] Penslar RB, Porter JP. 1993. Office for Human Research Protections (OHRP) IRB Guidebook. United States Department of Health and Human Services (HHS). Available: *http://www.hhs.gov/ohrp/archive/irb/irb_guidebook.htm* [accessed 15 February 2012].

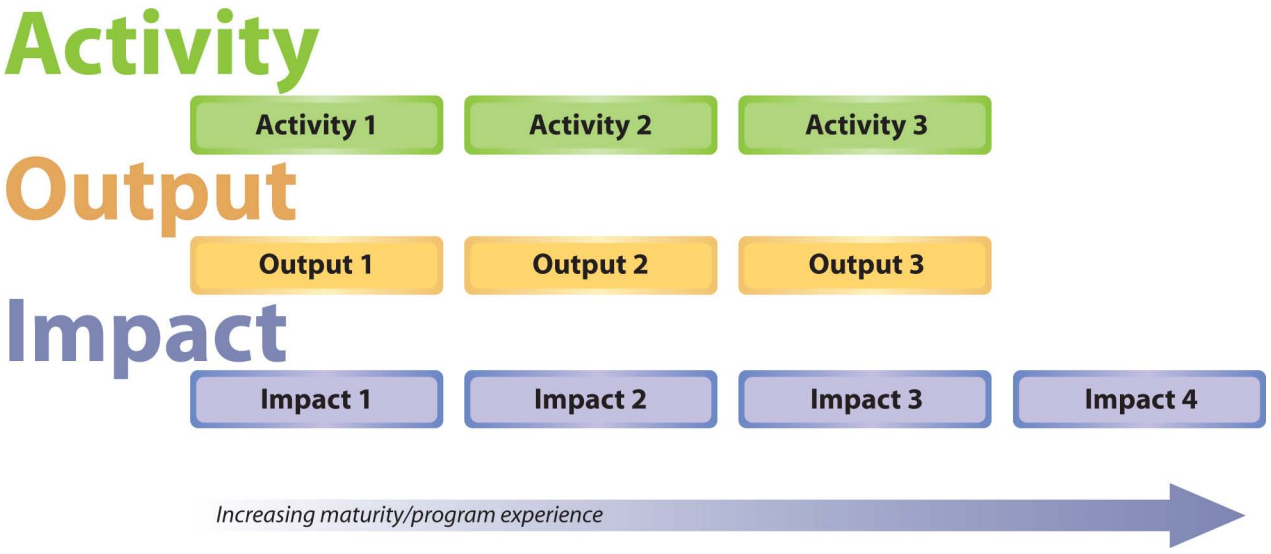
Logic Models

This Manual makes extensive use of logic models as an approach to developing metrics. A logic model "presents a plausible and sensible model of how the program will work under certain conditions to solve identified problems."[68] It is a framework for showing the relationship between the activities a project conducts and the ultimate impacts or outcomes it achieves. Logic models illustrate the key elements of a project, help identify the relationships between project activities and goals, and describe the intended impacts and how they can be measured. Perhaps most importantly, logic models are a tool for showing the cause-and-effect relationships between the project and its goals.[69]

There are many benefits of using a logic model. The process of developing program logic models may contribute to strategic planning by providing partners with a way to build consensus about a project's purpose and by identifying necessary resources. A completed logic model can be a useful tool to illustrate the project design and objectives for staff, partners, funders and decision-makers. The logic model can be used as a communication tool with both partners and parties external to the project. Finally, logic models can provide a framework for identifying metrics to measure project success as well as for identifying areas that need improvements. Such a framework can be used to develop an evaluation plan and provide feedback mechanisms for project leadership.

For simplicity (and to enable a greater focus on how to develop project metrics), the logic models described in Chapters 2 through 6 of this Manual have focused primarily on activities, outputs and impacts (Figure 7.1). However, logic models typically include several other components to further illustrate and describe various program processes and characteristics. In this section, we describe inputs, contextual factors and ultimate impacts, and we provide examples of how these elements may be useful for project planning and evaluation.

Figure 7.1 Format of the Logic Model Example Used in the PEPH Evaluation Metrics Manual



[68] McLaughlin JA, Jordan GB. 1999. Logic Models: A tool for telling your program's performance story. Eval Program Plann 22(1).

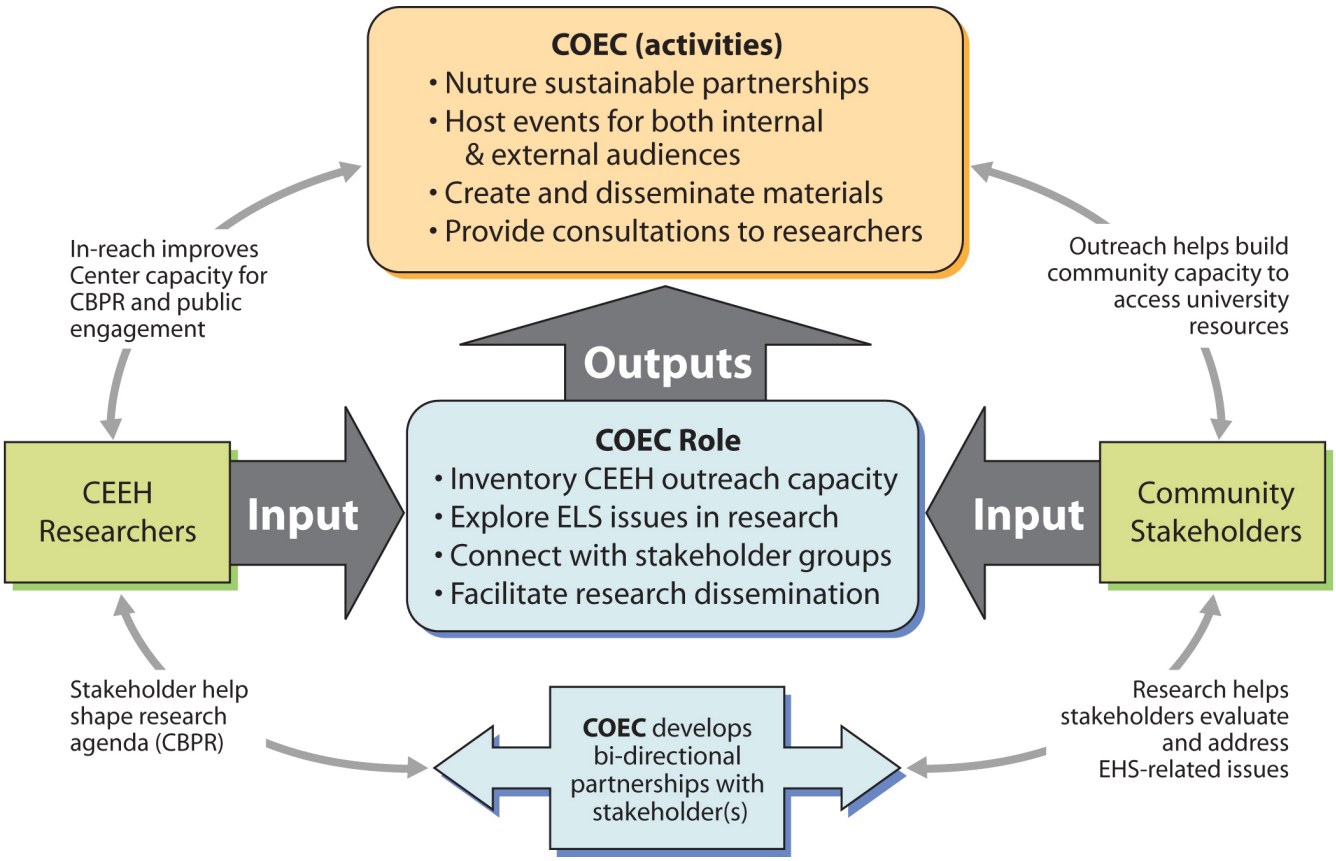
[69] Watson S. 2002. Learning from Logic Models in Out-of-School Time. Harvard Family Research Project. Available: http://www.hfrp.org/evaluation/the-evaluation-exchange/issue-archive/evaluating-out-of-school-time/logic-models-in-out-of-school-time [accessed 15 February 2012].

Inputs

Inputs encompass all of the assets available to partners to allow them to accomplish their project goals, and they include human, financial, organizational and community resources. Inputs can be tangible, such as a group of volunteers or grant funding, or intangible, such as a partnership. They can also be intellectual (ideas), material (equipment) and logistical (people's time). Lastly, inputs may include the major forces that influence the organization or program, such as the regulatory framework or political state of affairs. As an example, we provide the program logic model for the Community Outreach and Ethics Core (COEC) at the Center for Ecogenetics and Environmental Health (CEEH) at the University of Washington (see Figure 7.2).

In this example, environmental health researchers and community members are the human resource inputs. The model highlights the role that leveraging and capacity building can play in a PEPH project, demonstrating how leveraging community partners and CEEH researchers can lead to increased community and CEEH capacity. The methods outlined in Chapters 3 and 6 on leveraging and capacity building provide more information about assessing and gathering initial inputs, as well as building upon existing resources.

Figure 7.2 Logic Model of the Community Outreach and Ethics Core (COEC) at the University of Washington[70]



[70] Center for Ecogenetics and Environmental Health (CEEH) at the University of Washington. 2010. CEEH Outreach. Available: *http://depts.washington.edu/ceeh/community_ceeh.html* [accessed 16 December 2011].

Contextual factors

Contextual factors describe the economic, social and political environment that might influence the implementation or the impacts of the program and are beyond the control of the program staff. Examples of contextual factors include a disease outbreak, a storm that disrupts data collection, election results and state or federal budget reductions. While program staff cannot control contextual factors, they can anticipate, plan for and adapt to them.[71]

Ultimate impacts

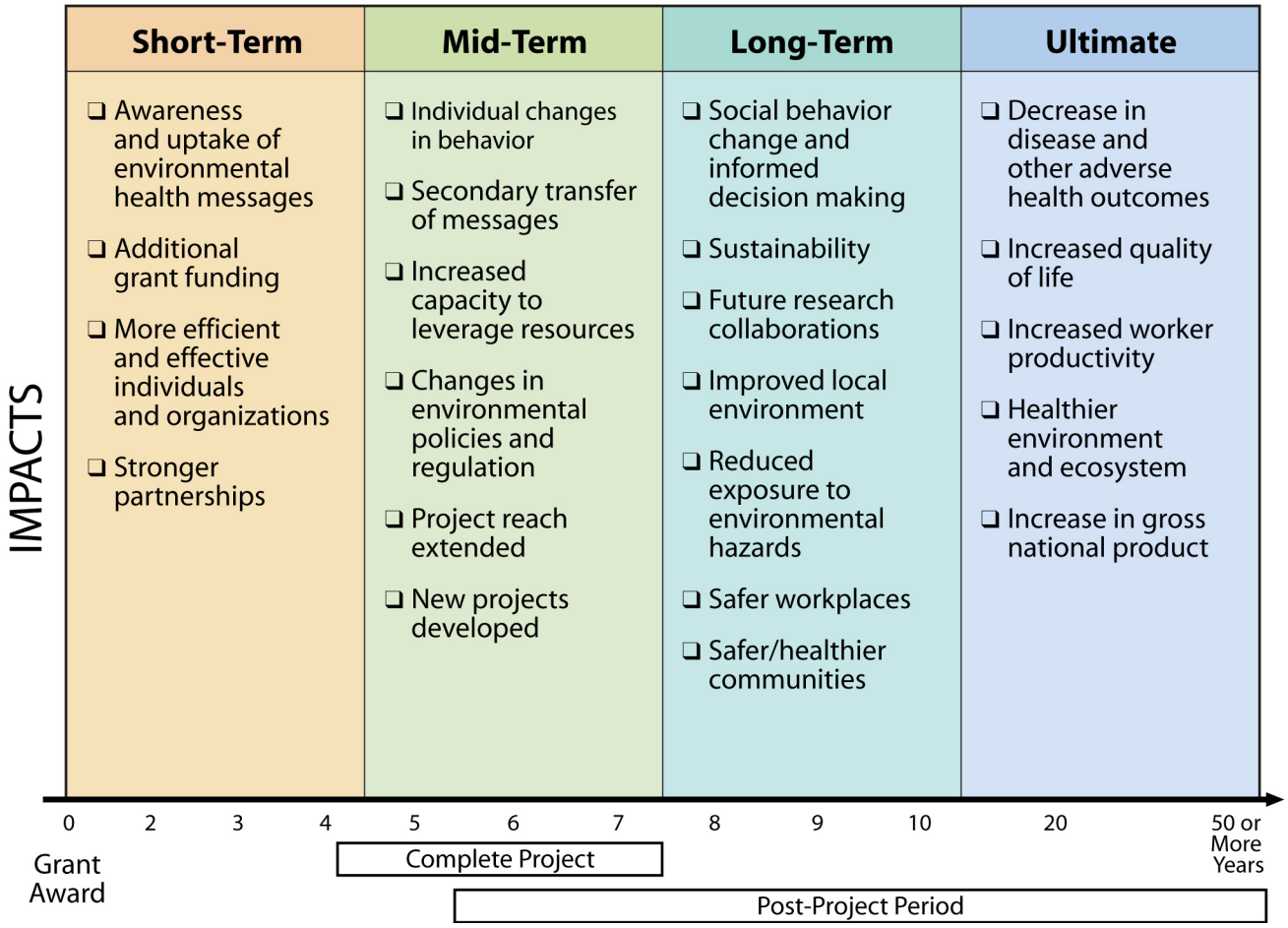
Ultimate impacts refer to the future societal change grantees hope to achieve with a project. These are sometimes called long-term outcomes or impacts and can appear decades after project activities have begun. They generally fall into two categories: 1) improved human health and well-being and 2) benefit to the economy. The ultimate impacts on human health and well-being of a PEPH project could include a decrease in disease or other adverse health outcomes associated with environmental health agents. A decrease in adverse environmental health hazards and illness may ultimately benefit the economy through a reduction in work and school absences, improvement in worker productivity and a decline in health care costs. The target population could also reap the ecological benefits of a healthier environment and ecosystem. Figure 7.3 shows possible impacts stemming from a PEPH project grant in the short-term, mid-term, long-term and ultimate time frames.

Arrows

Other important features of logic models that are not included in our simplified version are the arrows that show the interactions between the various components of the logic model. The direction and flow of the arrows can be adapted to reflect the unique characteristics of each program.

[71] Centers for Disease Control and Prevention (CDC). 1999. Framework for program evaluation in public health. MMWR 48(RR-11). Available: *http://www.cdc.gov/mmwr/pdf/rr/rr4811.pdf* [accessed 16 December 2011].

Figure 7.3 Project Evaluation Timeline Showing Examples of Short-term, Mid-term, Long-term and Ultimate Impacts[72]



Use of Logic Models at NIEHS

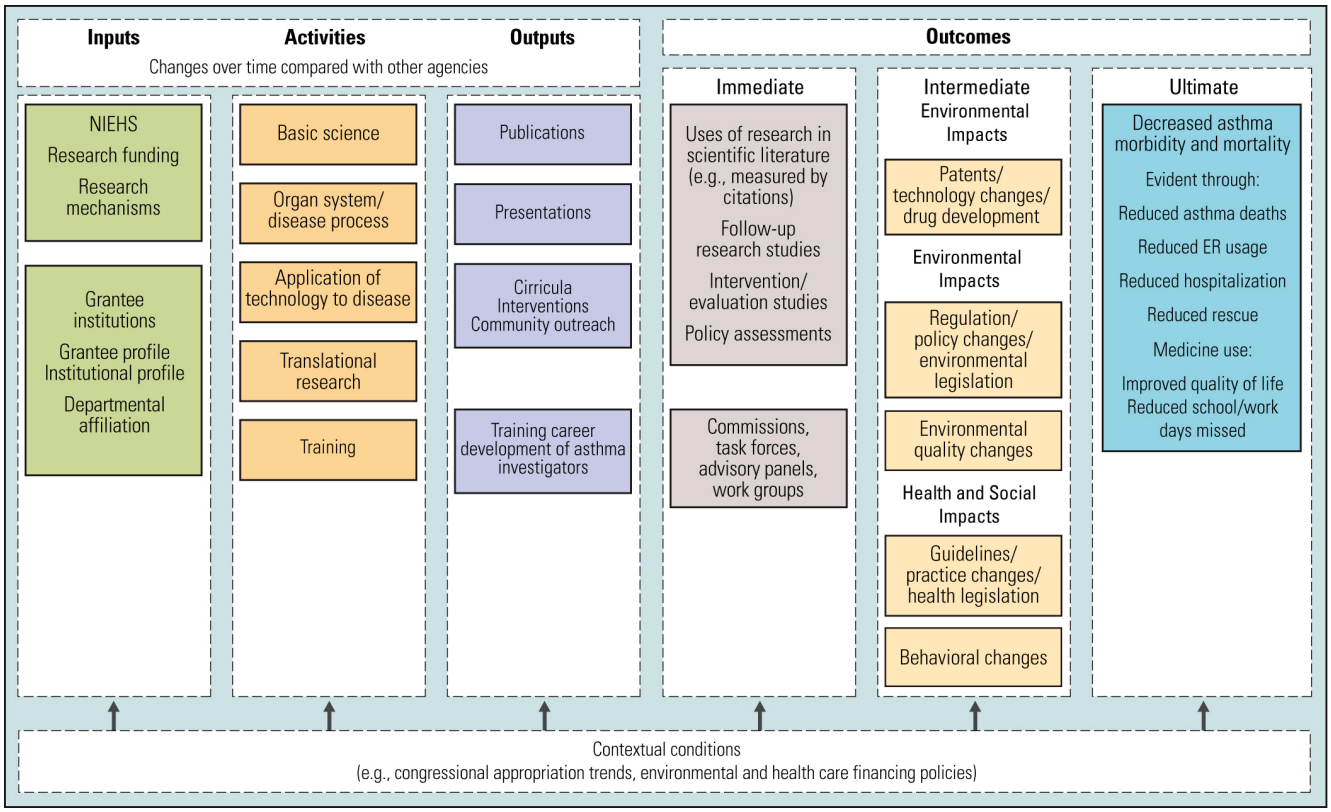
In recent years, NIEHS has been using logic models to examine the long-term impacts of its investments in research grants. For example, a major effort to evaluate the long-term impacts of the NIEHS Asthma research portfolio was conducted from 1975 to 2005.[73,74,75] A complex logic model was developed for this purpose (Figure 7.4). This model illustrates the link between NIEHS-funded activities and outputs, with the intended ultimate outcome of decreased asthma morbidity and mortality. It also highlights immediate outcomes, such as task forces, and intermediate outcomes, such as drug development. Contextual conditions (e.g. healthcare financing policies) are presented across the bottom of the figure as possible influences on inputs, activities, outputs and outcomes.

[72] Adapted from, Ruegg, R. 1999. Assessment of the ATP. In: The Advanced Technology Program, Challenges and Opportunity. Washington, DC: National Academy Press. 19.

[73] Engel-Cox J, Van Houten B, Phelps J, Rose S. 2008. Conceptual model of comprehensive research metrics for improved human health and environment. *Environ Health Perspect* 116(5): 583-92. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2367676/?tool=pubmed [accessed 16 December 2011].

[74] Orians CE, Abed J, Drew CH, Rose SW, Cohen JH, Phelps J. 2009. Scientific and public health impacts of the NIEHS extramural asthma research program: Insights from primary data. Res Evaluat 18(5):375-385.

Figure 7.4 NIEHS Asthma Research Portfolio Logic Model[75]



The NIEHS Worker Education and Training Program (NIEHS WETP) also uses a logic model to describe its program, including the outputs and impacts the project expects to see from grantees (Figure 7.5). The WETP program provides occupational safety and health training to workers who handle hazardous materials or respond to emergency releases of hazardous materials. There are five training programs:

- The Hazardous Waste Worker Training Program provides model occupational safety and health training for workers who are or may be engaged in activities related to hazardous waste removal, containment or chemical emergency response.
- The Minority Worker Training Program focuses on delivering comprehensive training to disadvantaged inner city young adults in order to prepare them for employment in the fields of environmental restoration and hazardous materials.
- The NIEHS/Department of Energy (DOE) Nuclear Worker Training Program trains workers engaged in environmental restoration, waste treatment and emergency response at sites in the DOE's nuclear weapons complex.

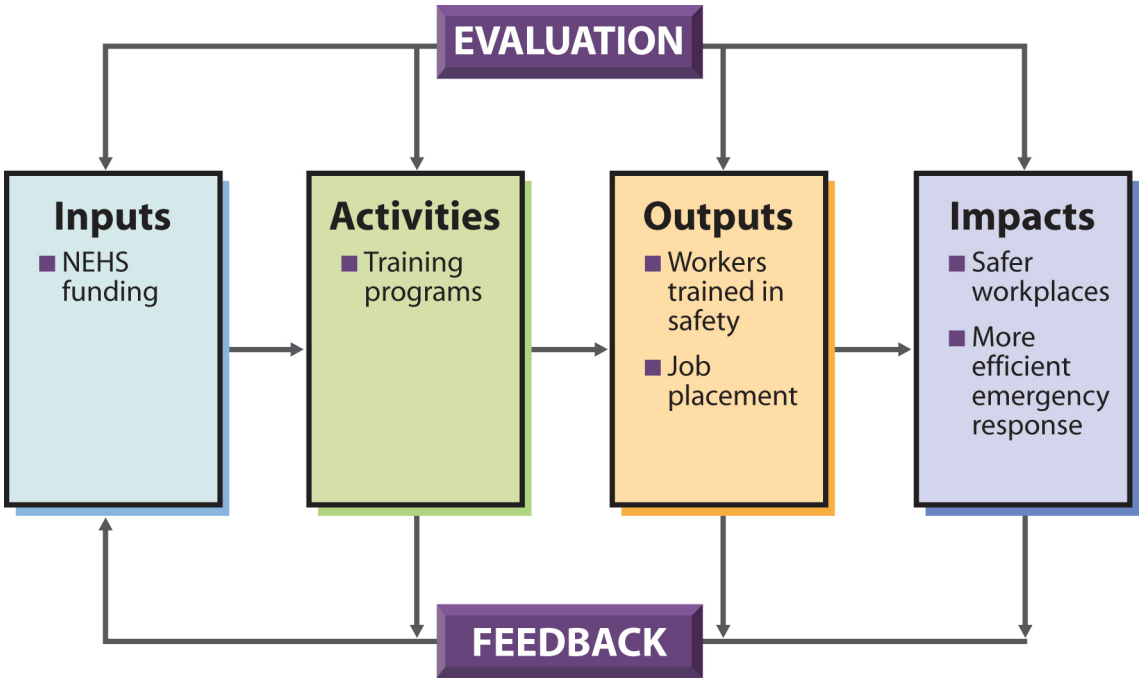
[75] Liebow E, Phelps J, Van Houten B, Rose S, Orians C, Cohen J, et al. 2009. Toward the assessment of scientific and public health impacts of the National Institute of Environmental Health Sciences Extramural Asthma Research Program using available data. *Environ Health Perspect* 117(7). Available: http://www.ehponline.org/ambra-doi-resolver/10.1289/ehp.0800476 [accessed 15 February 2012].

- The Hazmat Disaster Preparedness Training Program enhances the safety and health training of current hazardous materials workers and chemical responders and augments prevention and preparedness efforts in a wide variety of high-risk settings.
- The Advanced Training Technology Program focuses on the development of training products for health and safety training for hazardous materials workers, emergency responders and skilled support personnel.

Nonprofit training centers perform the actual training with the help of a NIEHS grant, but NIEHS WETP evaluates each of the five overall programs. To assess their progress, the WETP program conducts annual evaluations focusing on training and job placement as key indicators of success, and it publishes the results. NIEHS uses evaluation to ensure that the independent training centers are achieving their intended outputs and impacts.

To see examples of evaluation reports from WETP training programs, visit:
http://tools.niehs.nih.gov/wetp/index.cfm?id=92

Figure 7.5 NIEHS WETP's Logic Model of its Evaluation of Inputs, Activities, Outputs and Impacts



Types of Evaluation

Grantees may conduct evaluations for a variety of reasons. Different types of evaluations can be used to answer different types of questions.[76] The descriptions below provide an overview of four of the primary types of evaluations.

PROCESS EVALUATION

This form of evaluation assesses the extent to which a program is operating as it was intended. It typically assesses program activities' conformance to statutory and regulatory requirements, program design and professional standards or customer expectations.

OUTCOME EVALUATION

This form of evaluation assesses the extent to which a program achieves its outcome-oriented objectives. It focuses on outputs and outcomes (including unintended effects) to judge program effectiveness but may also assess program process to understand how outcomes are produced.

IMPACT EVALUATION

Impact evaluation is a form of outcome evaluation that assesses the net effect of a program by comparing program outcomes with an estimate of what would have happened in the absence of the program. This form of evaluation is employed to isolate the program's contribution to achievement of its objectives when external factors are known to influence the program's outcomes.

COST-BENEFIT AND COST-EFFECTIVENESS ANALYSES

These analyses compare a program's outputs or outcomes with the costs (resources expended) to produce them. When applied to existing programs, they are also considered a form of program evaluation. Cost-effectiveness analysis assesses the cost of meeting a single goal or objective and can be used to identify the least costly alternative for meeting that goal. Cost-benefit analysis aims to identify all relevant costs and benefits, usually expressed in dollar terms.

[76] U.S. Government Accountability Office (GAO). 2011. Performance measurement and evaluation. GAO 11-646SP. Available: *http://www.gao.gov/new.items/d11646sp.pdf* [accessed 15 February 2012].

Planning an Evaluation

After identifying the intended activities, outputs and impacts of a program, grantees should have the information necessary to begin planning an effective program evaluation. An evaluation plan provides a formal opportunity for grantees to document the steps they will take to conduct a program evaluation.

An evaluation plan typically includes descriptions of the following:

- Purpose of program
- Partner assessment
- Evaluation goals
- Evaluation questions
- Data collection plans
- Data analysis plans
- Dissemination and reporting activities
- Other evaluation products
- Timeline and budget
- Staff responsible for each evaluation activity

In the next section, we provide more details about data collection, data analysis, and reporting and dissemination.

For sample evaluation plans, check out the following sources:
http://managementhelp.org/evaluatn/chklist.htm
http://www.epa.gov/evaluate/pdf/evalworksheet.pdf
http://www.wmich.edu/evalctr/wp-content/uploads/2010/05/plans_operations1.pdf

Data collection

Data can be categorized as either qualitative or quantitative. Qualitative data are descriptions of the characteristics of that which is being analyzed. Qualitative data are often collected in open-ended questions, feedback surveys or summary reports. Qualitative data provide valuable and insightful data but can be difficult to compare, reproduce and generalize. Quantitative data are numerical or statistical values used to express the quantities of a variable. This type of data is relatively easy to store and manage and can be generalized and reproduced, but it usually fails to provide a complete picture of a program.

A mixed-methods approach that combines quantitative and qualitative data can provide a more complete picture of a program.

When conducting an evaluation, partners can use different types of evidence: logs and documents directly associated with their project, data gathered from community members or other participants, research from external sources and environmental and health data. Some examples of data sources for each type of evidence are presented below (see Table 7.1). Data collection can be performed by partners or obtained from external sources. For example, partners can personally gather health data on the incidence of a particular disease in their community or obtain external government statistics from the Centers for Disease Control and Prevention (CDC). Likewise, community members can conduct their own environmental study or use data collected by other organizations, such as the Environmental Protection Agency (EPA), National Institute for Occupational Safety and Health (NIOSH), etc.

Data analysis

Data analysis plans provide an opportunity for grantees to think about what methods they are going to use to answer the evaluation questions. Content analysis, social network analysis and bibliometric analyses are methods grantees can use to organize and understand qualitative data. Often basic spreadsheet and word-processing software are all that is needed to conduct qualitative analyses. However, specialized qualitative analysis software such as Atlas.ti and NVivo are available to help organize and code data. Qualitative data can be analyzed on a case-study basis where each subject is analyzed and understood on its own, or by grouping similar "subjects" together.

For a list of currently available software and links to developer websites, see American Evaluation Association, "Qualitative Software," **http://www.eval.org/Resources/QDA.asp**.

See also U.S. General Accounting Office (GAO), "Quantitative Data Analysis: *An Introduction," Report to Program Evaluation and Methodology Division*, May 1992, **http://www.gao.gov/special.pubs/pe10111.pdf**.

Table 7.1 Examples of Data Sources for Evaluations

Type of Data Sources	Examples of Data Sources
Project Logs	Lists of partners/attendees at meetings Activity reports Meeting summaries Video and tape recordings
Project Documents	Study questions Logic model Project plan Quarterly/annual reports Governance agreements/documents Budget documents Educational products from project
Data Collected during Project	Diaries or field notes Forms Surveys Interviews Anecdotal evidence/stories Observations Behavioral data
Research from External Sources	Official records Letters Newspaper accounts Published data Ethnographies Oral histories
Environmental	Exposure to environmental toxins Water quality data Air quality data
Health	Incidence/prevalence of diseases or injuries Health-related behavior, knowledge and skills

Quantitative analysis describes any method for organizing or understanding numerical data. Examples of quantitative analysis methods include:

- Descriptive statistics
- Linear models
- Correlations and regressions
- Return on investments

As with qualitative analysis, a basic spreadsheet program is all that is needed to answer most quantitative evaluation questions. However, software such as SAS, SPSS, and STATA are useful for conducting more complex statistical analyses.

Metrics in Action 7.2 provides an example from a PEPH program that incorporates both qualitative and quantitative data analysis.

Metrics in Action 7.2: The **University of Texas Medical Branch-Galveston (UTMB) Center to Eliminate Health Disparities**, in conjunction with NIEHS COEC and the Community In-Power Development Association, Inc. (CIDA) of Port Arthur, Texas, uses data analysis to develop and apply a cumulative risk framework to address the community's environmental justice concerns. The Center partners initially focused on merging environmental and social determinants of health into a single, integrated assessment. Researchers used the following data sources:

- Census data
- Aggregated Texas Department of State Health Services health data
- EPA Toxics Release Inventory and Texas Commission on Environmental Quality monitoring data
- Occupational Safety and Health Administration (OSHA) safety data
- Documentation of industrial accidents, explosions and flaring

Children playing in municipal park next to chemical refineries in Port Arthur, Texas. *Photo by H. Kelley*

- Results of community symptom surveys, community interviews, focus groups and arts-based popular education and communication interventions
- Maps of key indicators of environmental and social risk using a geographic information system and community-mapping workshops

Initial results show that multiple stressors and health disparities disproportionately affect West Port Arthur, and residents in this area are exposed to significant cumulative risk burdens. The use of both qualitative and quantitative data analysis allowed the Center researchers to accurately map a range of indicators of the community's overall risk burden.

For more information on the Center to Eliminate Health Disparities, visit: http://www.utmb.edu/cehd/.

Reporting and dissemination

Evaluation findings can support actions to improve PEPH projects by identifying strengths and weaknesses or suggesting modifications to underlying organizational systems. Demonstrating the effectiveness of a PEPH project via an evaluation can result in improved accountability, quality control or increased project scope or funding. Evaluations can lead to the generation of new and enhanced knowledge and theories specifically for environmental public health, or more broadly for human and organizational development. Finally, evaluation findings can inspire policy changes affecting a population far beyond the original scope of the PEPH project.

Once the project has been evaluated, it is important to consider what to do with the findings. PEPH partners might want to ensure that the use of the evaluation is consistent with the original purpose of the project.[77] For example, if partners choose to evaluate the impact an education program has had on different community groups, they could share their findings not just with the funding agency, but also with those very community groups. This could end up furthering the original program goal of education while disseminating the results. Sharing results could also lead to the improvement of projects other than the one under direct evaluation.[78] In planning to share evaluation results, partners can ask:

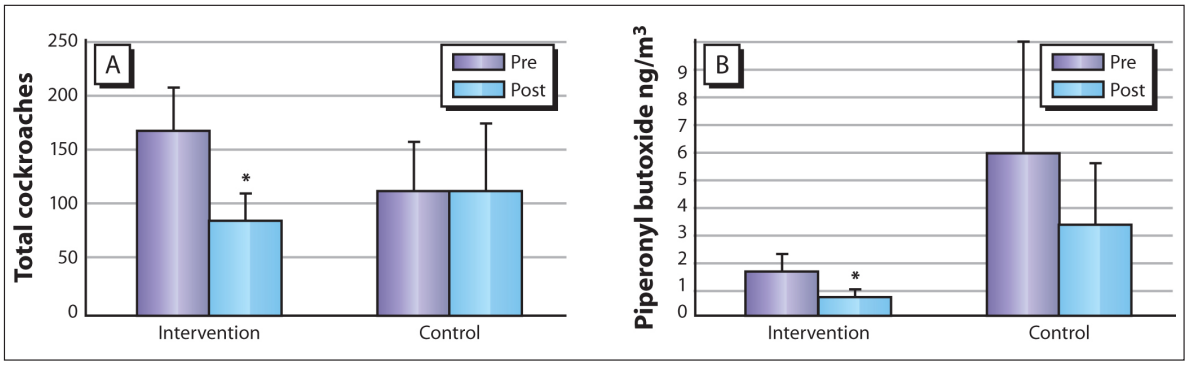
- What did partners learn from the project evaluation?
- Who might be interested in these results?
- How should the project's accomplishments be reported?
- How can the partners use these results to improve the program? Do the partners need to change project activities and objectives?
- How can the partners use the results to secure additional funding?
- How would the partners use these results to assess impacts over a longer time frame?
- What cultural or confidentiality issues do the partners need to address?

[77] Frechtling J. 2002. The 2002 User Friendly Handbook for Project Evaluation. Arlington, VA: The National Science Foundation, Division of Research, Evaluation, and Communication. 71-2. Available: *http://www.nsf.gov/pubs/2002/nsf02057/nsf02057.pdf* [accessed 16 December 2011].

[78] Barnes, H, Jordan G. 2006. EERE Guide for Managing General Program Evaluation Studies: Getting the Information You Need. Office of Energy Efficiency and Renewable Energy (EERE), Office of Planning, Budget and Analysis. 57. Available: *http://www1.eere.energy.gov/ba/pba/pdfs/evaluation_mgmt_guide_final_2006.pdf* [accessed 16 December 2011].

Metrics in Action 7.3: Researchers at the **Columbia Center for Children's Environmental Health (CCCEH)** evaluated the effectiveness of using integrated pest management (IPM) to reduce both pest infestation (cockroaches) and insecticide exposure after documenting widespread exposure to insecticides among pregnant inner-city women in Harlem, New York. The IPM program uses a variety of methods, including professional cleaning, sealing of pest entry points, application of low-toxicity pesticides and education. The evaluation revealed that pest levels significantly decreased in the IPM intervention households, but not in the control households. Likewise, levels of pyrethroid insecticides in indoor air samples were significantly lower in intervention households than in control households (Figure 7.6). Furthermore, researchers detected the presence of insecticides in blood samples of mothers in the control group, but not in the IPM intervention group. The evaluation successfully demonstrated the effectiveness of using IPM to reduce pest infestation and insecticide exposure during pregnancy. The researchers then published their results in *Environmental Health Perspectives* to disseminate the findings to the academic community. They also educated the residents of Harlem on the risks of pesticides for pregnant women and ways to mitigate these risks using IPM. By reporting the results of their evaluation and performing outreach, the CCCEH shared its best practices with others.

Figure 7.6 Cockroach Infestation Levels (left: A) and the Use of Piperonyl Butoxide (right: B) in 2-Week Integrated Air Samples



For more information about the CCCEH, visit: http://www.cumc.columbia.edu/dept/mailman/ccceh.

Additional Resources

The intent of this chapter is to provide a broad overview of evaluation practice, but it is by no means comprehensive. See Appendix 4 for additional resources and publications concerning:

- General program evaluations
- Environmental health and health program evaluations
- Logic models
- Evaluation tools
- Process evaluations
- Outcome evaluations
- Online databases