

What is program evaluation?

A beginners guide

This guide is available as a set of smaller pamphlets here
<http://gsociology.icaap.org/methods/BasicguidesHandouts.html>

Produced by

Gene Shackman, Ph.D.
The Global Social Change Research Project
<http://gsociology.icaap.org/>
Free Resources for Methods in Program Evaluation
<http://gsociology.icaap.org/methods>

What is evaluation

When beginning an evaluation, program people will often want the answer to this question:

- Does the program work? And how can it be improved?

However, there are many equally important questions

- Is the program worthwhile?
- Are there alternatives that would be better?
- Are there unintended consequences?
- Are the program goals appropriate and useful?

This handout focuses on the first of these issues: how program evaluation can contribute to improving program services.

Evaluations, and those who request them, may often benefit, though, from a consideration of these other questions.

An evaluation can help a program improve their services, but can also help ensure that the program is delivering the **right** services.

See this resource for additional information:

Developing a Concept of Extension Program Evaluation
<http://learningstore.uwex.edu/Designing-Evaluations-C238.aspx>
Mohammad Douglah, University of Wisconsin, Cooperative Extension. 1998.

What is evaluation

In this pamphlet, we describe program evaluations used to improve program services, using the definition below to organize our knowledge.

Program evaluation is:

“...the systematic assessment of the operation and/or outcomes of a program or policy, compared to a set of explicit or implicit standards as a means of contributing to the improvement of the program or policy...”*

* Carol Weiss, quoted in
Introduction to Program Evaluation
<http://www.cdc.gov/tb/programs/Evaluation/Default.htm>
See slide #4
Quote used in this pamphlet by permission.

What is evaluation

One main goal of program evaluation is:

“contributing to the improvement of the program or policy”

This handout describes some of the ways that program evaluation can help improve program services, in particular, briefly describing:

- Planning the evaluation
- Determining the evaluation questions
- Answering evaluation questions: evaluation methods

What is evaluation

Lets start with this part of evaluation:

“...the systematic assessment”

An evaluation is a systematic assessment. Evaluations should follow a systematic and mutually agreed on plan. Plans will typically include the following:

- Determining the goal of the evaluation: What is the evaluation question, what is the evaluation to find out.
- How will the evaluation answer the question: What methods will be used.
- Making the results useful, how will the results be reported so that they can be used by the organization to make improvements.

Additional resources about planning evaluations:

Centers for Disease Control and Prevention. Framework for Program Evaluation in Public Health. MMWR 1999;48(No. RR-11). <http://www.cdc.gov/eval/framework.htm>

The Planning-Evaluation Cycle
<http://www.socialresearchmethods.net/kb/pecycle.php>
In Research Methods Knowledge Base, by William M.K. Trochim
<http://www.socialresearchmethods.net/kb/index.php>

What is evaluation

The first part of the evaluation is to determine the question.

“assessment of the operation and/or outcomes of a program or policy”

Evaluations can generally answer two types of questions:

1. What is the outcome of the program? Did the program have any impact, was there any improvement in people's lives?
2. How did the program get to that outcome? Did the program have some set of procedures? Were these procedures followed, were the procedures reasonable, was there a better way to get to the outcomes?

Additional resources:

Approaching An Evaluation-- Ten Issues to Consider
Brad Rose Consulting, Inc.
http://www.bradroseconsulting.com/Approaching_an_Evaluation.html

What is evaluation

Back to determining the evaluation question.

One way to do this is for the evaluator and program people to develop a very good description of:

- what the outcomes should be,
- how the program will get there, and
- why the program leads to the outcome.

This description helps to identify how the program should lead to the outcome, why the program activities should lead to the outcomes, and where to evaluate the program to check whether it does.

This method is called a *program theory*.

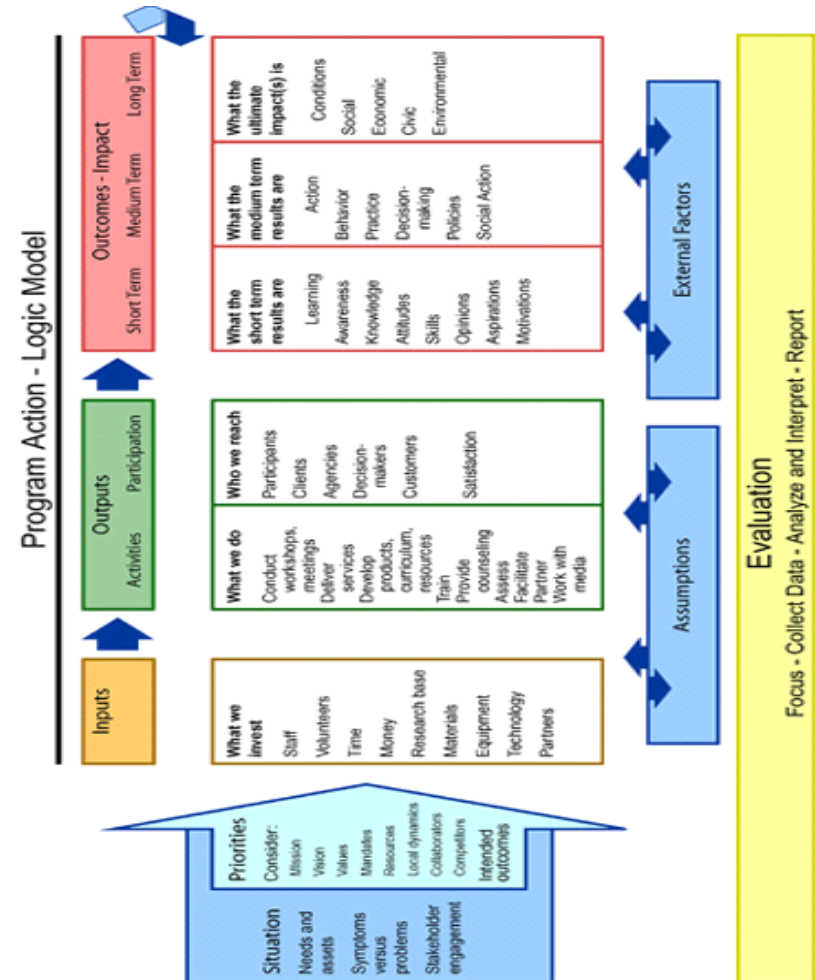
“A program theory explains how and why a program is supposed to work. ... It provides a logical and reasonable description of why the things you do – your program activities – should lead to the intended results or benefits.”

From Program Evaluation Tip Sheets from Wilder Research, Issue 4, October 2005 - Program Theory. http://www.ojp.state.mn.us/grants/Program_Evaluation/

A useful tool to help work with the program theory is a *logic model*, which visually shows the program theory, how all the program goals, activities, and expected outcomes link together.

What is evaluation

Logic model example:



from
 Logic Model, University of Wisconsin Extension, Program
 Development and Evaluation.
<http://www.uwex.edu/ces/pdande/evaluation/evallogicmodel.html>

What is evaluation

Use the program theory or logic model to come up with evaluation questions

- Does the program have a positive outcome?
- Are people satisfied?
- How could the program be improved?
- How well is the program working?
- Is the program working the way it was intended to work?

Additional Resources

Developing Evaluation Questions
Mid-Continent Comprehensive Center
<http://www.mc3edsupport.org/community/knowledgebases/developing-evaluation-questions-820.html>

Developing Process Evaluation Questions. At the National Center for Chronic Disease Prevention and Health Promotion. Healthy Youth. Program Evaluation Resources
<http://www.cdc.gov/healthyyouth/evaluation/resources.htm>

What is evaluation

However, there are limits to program theory and logic models:

- Models are linear, programs are complex, interactive
- Models are static, programs may change over time.
- Models may not take unexpected consequences into account
- Models may not account for conflict, power, control issues
- Theory or model assumes the model is correct.

Use program theory and logic models, but be flexible, and open to change and feedback. Review and revise them often, as necessary.

Additional Resources about logic models.

Usable Knowledge's Interactive logic model tutorial
http://www.usablellc.net/html/links_we_like.html#logic_models

Program logic - an introduction
from Audience Dialogue
<http://www.audiencedialogue.net/proglog.html>

From: Logic Model Basics. At the National Center for Chronic Disease Prevention and Health Promotion. Healthy Youth. Program Evaluation Resources
<http://www.cdc.gov/healthyyouth/evaluation/resources.htm>

A Guide on Logic Model Development for CDCs Prevention Research Centers (Sundra, Scherer, and Anderson)
<http://www.ojp.usdoj.gov/BJA/evaluation/guide/pe4.htm>

What is evaluation

Randomly or systematically choosing people to respond to surveys means using some defined method to select people. For example:

- Randomly choosing – generate a list of random numbers, assign each person a random number, sort the people by the random number and take the people listed first. They were put on top of the list randomly.
- Systematic selection – a typical method is to start with the 5th person and then select every 7th person after that. The numbers, the 5th and the 7th are also chosen randomly.
- Randomly select locations to be in the sample, and then survey everyone in that location.

Random or systematic selection means that the group of people you select are more likely to be similar to your clients, in general. You aren't excluding any particular groups, or including only certain groups. You are avoiding bias, in sampling terms.

If you **do** use random or systematic selection, then most likely you **can** use the results of your survey to make conclusions about your clients.

If you **don't** use random or systematic selection, you can **NOT** use the results of your survey to make conclusions about your clients population. That is, you cannot generalize from your study to your client population. You can only say “The people who took this survey said ...”

What is evaluation

Additional Resources about surveys

What is a Survey, by Fritz Scheuren
<http://www.whatisasurvey.info/>

Surveys and You.
From the Council of American Survey Research Organizations
<http://www.casro.org/survandyou.cfm>

Sampling
<http://www.socialresearchmethods.net/kb/sampling.php>
In Research Methods Knowledge Base, by William M.K. Trochim
<http://www.socialresearchmethods.net/kb/index.php>

Sampling
<http://www2.chass.ncsu.edu/garson/pa765/sampling.htm>
in Statnotes: Topics in Multivariate Analysis, by G. David Garson
<http://www2.chass.ncsu.edu/garson/pa765/statnote.htm>

What is evaluation

Analysis of administrative data is just using statistical analysis on program data that is already collected.

Administrative data has advantages:

- No new data collection is required
- Many databases are relatively large
- Data may be available electronically

and disadvantages:

- Data were gathered for another purpose, so may not have necessary variables.
- In all administrative data sets, some fields are likely to be more accurate than others.

Using Administrative Data To Monitor Access, Identify Disparities, and Assess Performance of the Safety Net. By John Billings, J.D. Tools for Monitoring the Health Care Safety Net. Agency for Healthcare Research and Quality.
<http://www.ahrq.gov/data/safetynet/billings.htm>

Additional Resources

Data collection: Types of data collection – Administrative Data. Statistics Canada.
<http://www.statcan.ca/english/edu/power/ch2/types/types.htm#administrative>

What is evaluation

Focus groups are structured discussions among small groups of people.

Generally, a facilitator leads a group of 8-10 people in a discussion about selected topics with planned questions, while allowing for interesting, new or unplanned follow up questions.

Typical focus group questions are like these:

- What is your overall impression of the program?
- What are the things you like or dislike about the program?
- What have you gained in this program?
- If you have not noticed any changes in yourself, what do you think are the reasons?

From: Qualitative Evaluation of the Project P.A.T.H.S. Based on the Perceptions of the Program Participants. Shek, Daniel T.L., Lee, Tak Yan, Siu, Andrew, Lam, Ching Man. The Scientific World Journal. November 2006, 1, 2254–2264
<http://www.thescientificworld.co.uk/TSW/main/home.asp?ocr=1&jid=141>

Additional Resources about focus groups

Basics of Conducting Focus Groups
Carter McNamara, MBA, PhD, Authenticity Consulting, LLC
<http://www.managementhelp.org/evaluatn/focusgrp.htm>

Focus Groups.
From the National Park Service Northeast Region
http://www.nps.gov/nero/rtcatoolbox/gatinfo_focus.htm

What is evaluation

Key informant interviews are qualitative, in-depth interviews of 15 to 35 people selected for their first-hand knowledge about a topic of interest.

Conducting Key Informant Interviews. Performance Monitoring and Evaluation. USAID Center for Development Information and Evaluation.
http://www.usaid.gov/pubs/usaid_eval/

Key informant interviews also include a planned set of questions on the topics of interest.

Key informant interviews are useful when candid information about sensitive topics is needed. Group discussions may inhibit people from giving candid feedback.

Interviews should include a very diverse range of people.

Additional Resources

Key Informant Interviews
University of Illinois Extension
<http://ppa.aces.uiuc.edu/KeyInform.htm>

What is evaluation

Observations are methods that yield a systematic description of events or behaviors in the social setting chosen for study.

Observation methods can be highly structured, for example:

Systematic Social Observation - a field research method in which teams of researchers observe the object of study in its natural setting. Researchers record events as they see and hear them and do not rely upon others to describe or interpret events. The researchers follow well-specified procedures that can be duplicated.

Systematic Observation of Public Police: Applying Field Research Methods to Policy Issues. Stephen D. Mastrofski, Roger B. Parks, Albert J. Reiss, Jr., Robert E. Worden, Christina DeJong, Jeffrey B. Snipes, William Terrill. National Institute of Justice, December 1998.
<http://www.ojp.usdoj.gov/nij/pubs-sum/172859.htm>

Observations can also be unstructured, for example, participant observation, or taking an active part in group activities.

The premise underlying participant observation is that the researcher becomes a more effective observer by taking an active role in the performance of regular activities. In other words, knowledge gained through doing is of a higher quality than what is obtained only through observation. In many cases, involvement with ordinary chores will not only enhance the researcher's understanding of the processes, techniques, and words associated with these activities, but will also result in better rapport with informants.

Documenting Maritime Folklife: An Introductory Guide
Part 2: How to Document. Participant Observation
American Folklife Center. Library of Congress.
<http://www.loc.gov/folklife/maritime/twopo.html>

What is evaluation

Focus groups, interviews and observation are *qualitative* research methods, that is, methods that are less likely to rely on statistical analysis.

Advantages

- Useful to help figure out major program problems that cannot be explained by more formal methods of analysis.
- The evaluator may see things that participants and staff may not see.
- The evaluator can learn about things which participants or staff may be unwilling to reveal in more formal methods
- Useful when it's not clear what the program problems might be.
- Useful to give good ideas of what topics program participants and staff think are important.
- Useful in developing surveys, in determining what questions or issues are important to include.
- Useful when a main purpose is to generate recommendations
- Useful when quantitative data collected through other methods need to be interpreted.

What is evaluation

Disadvantages

- The evaluator's subjective views can introduce error.
- The focus of the evaluator is only on what is observed at one time in one place.
- Information from observations/ interviews/ groups can be time consuming and difficult to interpret.
- Focus groups could be dominated by one individual and their point of view.
- Generally, information from focus groups, interviews, and observations CANNOT be used to describe the client population.

What is evaluation

Advantages and disadvantages of focus groups, observations and interviews **quoted from:**

The Handbook for Evaluating HIV Education - Booklet 9
Evaluation of HIV Prevention Programs Using Qualitative Methods
http://www.cdc.gov/HealthyYouth/publications/hiv_handbook/index.htm

Conducting Focus Group Interviews
USAID's Center for Development Information and Evaluation
http://www.usaid.gov/pubs/usaid_eval/

Conducting Key Informant Interviews. Performance Monitoring and Evaluation. USAID Center for Development Information and Evaluation. http://www.usaid.gov/pubs/usaid_eval/

Additional Resources:

Ethnography, Observational Research, and Narrative Inquiry:
Commentary
- Advantages of Qualitative Observational Research
- Disadvantages of Qualitative Observational Research
<http://writing.colostate.edu/guides/research/observe/pop2d.cfm>

Strengths: Data Collection Methods
Washington State Library, Connecting Learners to Libraries, 2006
Retreat
http://www.sos.wa.gov/library/libraries/projects/connecting/retreat_2006.aspx

Different Methods of Collecting Information
in What's the Best Way to Collect My Information?
<http://www2.ed.gov/pubs/EdTechGuide/whatbest.html>

What is evaluation

Did the program have an effect?

The ultimate goal of a program is to improve people's lives. How do you know whether it did?

One commonly used way to find out whether the program improved people's lives is to ask whether the program **caused the outcome**. If the program caused the outcome, then one could argue that the program improved people's lives.

On the other hand, if the program **did not cause** the outcome, then one would argue that, since the program did not cause the outcome then the program did not improve people's lives.

How to figure this out?

Determining whether a program caused the outcome is one of the most difficult problems in evaluation, and not everyone agrees on how to do it. Some say that randomized experiments are the best way to establish causality. Others advocate in-depth case studies as best. The approach you take depends on how the evaluation will be used, who it is for, what the evaluation users will accept as credible evidence of causality, what resources are available for the evaluation, and how important it is to establish causality with considerable confidence. (This paragraph suggested by Michael Quinn Patton.).

What is evaluation

There are three approaches frequently used to establishing whether a program causes an outcome.

- comparison groups – comparing people in the program to people not in the program
- multiple evaluation methods – comparing results from several evaluations, each using different methods
- in depth case studies of programs and outcomes – showing that the links between what program participants experience and the outcomes attained are reasonable, empirically validated, and based on multiple sources and data. The linkages between program and outcomes are direct and observable. No alternative possible causes offer a better explanation.

The particular method that is used should reflect a careful discussion and understanding of the pros and cons of each method, and agreement among all parties involved.

What is evaluation

Comparisons and cause:

Comparison groups and random assignment

The idea is this:

Randomly assign people to either be in the program (the 'treatment' group) or to be not in the program (the 'comparison' group).

Since people in the 'treatment group' were randomly assigned, then before the program the two groups of people should be pretty much the same.

Measure the treatment group after they have been on the program and compare them to people in the comparison group.

After the program, if the 'treatment' group people are better off than are the comparison group people, then the difference should be from being in the program, and so it is reasonable to argue that the program caused that outcome.

Additional Resources:

Why do social experiments? In Policy Hub, National School of Government's Magenta Book.

http://www.nationalschool.gov.uk/policyhub/evaluating_policy/magenta_book/chapter7.asp

What is evaluation

Advantages and disadvantages of random assignment to treatment and comparison groups.

Advantages:

- Results provide clearest demonstration of whether a program causes an outcome.
- Provides results that are easiest to explain.

Disadvantages:

- Often not practical to do. Can't randomly assign people to program or not program, and may be unethical to randomly assign someone to no treatment.
- Randomly assigning people to be in the program is not how programs really work, so results of the evaluation may not apply to the program as it really exists.
- Can't be applied to causes that operate over the long term or to programs that are very complex.
- Can tell whether a program caused outcome, but doesn't give much in depth information about why or how.
- People in treatment group know they are getting treatment so outcome may be due to knowledge, not to treatment.

This is a summary of points from:
(Munck and Jay Verkuilen) "Research Designs,"
<http://www-rcf.usc.edu/~munck/research/methods.html>

The last point is one of many from:
A Summative Evaluation of RCT Methodology: & An Alternative Approach to Causal Research. Michael Scriven. Journal of MultiDisciplinary Evaluation, Volume 5, Number 9. <http://jmde.com/>

What is evaluation

Comparison groups and non random assignment When random assignment is not possible, quasi-experimental design can be used. In this method, people "are not randomly assigned to groups but statistical controls are used instead."

Quasi-experimental designs. In Statnotes: Topics in Multivariate Analysis, by G. David Garson <http://www2.chass.ncsu.edu/garson/pa765/design.htm#quasi>

There are several versions of this approach:

- Comparing people already on the program to those who are not on the program. One example is to observe (O) people before they join the program or there is an intervention (X), then observe both groups after :

Pretest-posttest design

-Intervention group	O ₁	X	O ₂
-Comparison group	O ₁		O ₂

- Measuring the client many times before they join the program (or before a new intervention) and many times afterward, then compare before to after. One example is:

Time series design

-Intervention group	O ₁	O ₂	X	O ₃	O ₄
---------------------	----------------	----------------	---	----------------	----------------

- Combination of the two above

Time series design

-Intervention group	O ₁	O ₂	X	O ₃	O ₄
-Control group	O ₁	O ₂		O ₃	O ₄

What is evaluation

A major challenge to non random assignment approaches is that people on the program may start off being very different from the people not on the program.

That is, only some people choose to be on the program. Something made these people different and it may be the something which caused the better outcome, not the program.

One way to deal with this is to collect as much information as possible on characteristics of the people and program that relate to the program outcome (what the program is supposed to do), and use this information in statistical analysis to “control” for the differences between people on the program vs people not on the program.

The problem is that there may be differences, some critical, that are not observed, and for which the evaluator has no data.

Additional Resources:

AllPsych On Line. By Dr. Christopher L. Heffner
Section 5.3 Quasi-Experimental Design
<http://allpsych.com/researchmethods/quasiexperimentaldesign.html>

Quasi-experimental designs. In Statnotes: Topics in Multivariate Analysis, by G. David Garson
<http://www2.chass.ncsu.edu/garson/pa765/design.htm#quasi>

Diagrams on previous page from:
Measuring the Difference: Guide to Planning and Evaluating Health Information Outreach. Stage 4, Planning Evaluation. National Network of Libraries of Medicine
<http://nnlm.gov/evaluation/guide/>

What is evaluation

Multiple evaluation methods could support the idea that the program causes the outcome if different sources agree.

For example, collect information from:

- Program participants
- Program staff
- Community members
- Subject experts
- Published research and reports

Collect data through many methods, for example:

- Surveys
- Interviews
- Observations
- Program administrative data

If data from different sources don't agree, it doesn't necessarily mean the results from any of the sources are not valid. However, the more agreement there is from different sources, the more confident you can be about your conclusions.

Additional Resources:

An Introduction to Mixed Method Research. By Jennifer Byrne and Áine Humble. Atlantic Research Centre for Family-Work Issues. Mount Saint Vincent University.
<http://www.msvu.ca/ARCFamilyWork/publications.asp>

What is evaluation

An in depth case study can be used to demonstrate the connection between the intervention and the outcome.

An in-depth case study documents in detail what a group of participants experienced in a program and any ways in which they have changed so that the evaluator and users of the evaluation can make a judgment about the likelihood that the program led to the observed changes. For example, a group of chronic alcoholics go through a residential chemical dependency program. Their participation is fully documented. They return home maintaining their sobriety. They attribute their sobriety to the program as do their families, friends, and program staff. These multiple sources agree on the documented causal chain. The links between what they experienced and the outcomes attained are reasonable, empirically validated, and based on multiple sources and data. The linkages between program and outcomes are direct and observable. No alternative possible causes offer a better explanation. The preponderance of evidence points to a conclusion about causality. Establishing causality involves both data and reasoning about the findings. (This paragraph contributed by Michael Quinn Patton.)

What is evaluation

Random assignment is often seen as a very clear way to show whether a program causes an outcome. However, random assignment is often not practical or reasonable.

Other methods such as non random assignment, multiple evaluation methods, or in depth case studies are more practical and can be used to give reasonable arguments about whether a program caused an outcome.

However, these methods are less certain in establishing that the program is the cause of the outcome. There may be other things going on that are unknown and these other things might really be the cause of the outcome. It is more difficult for these other methods to rule out other possible causes, although the other methods can, again, establish reasonable likelihood. If a cause cannot be established, the evaluation can be used to describe what happened.

For example, the evaluation could say, “After the program, people were better off.” This doesn't necessarily mean it was the program that made the people better off, but the program may be one reasonable cause. Gathering more evidence, from multiple methods, driven by a very clear understanding of the program, can help determine the most reasonable explanation of causes of the outcomes.

What is evaluation

Additional Resources about design:

Steps in Program Evaluation

Gather Credible Evidence

<http://www.cdc.gov/eval/steps.htm#evidence>

Program Evaluation: A Variety of Rigorous Methods Can Help Identify Effective Interventions

GAO-10-30 November 23, 2009

<http://www.gao.gov/products/GAO-10-30>

What Counts as Credible Evidence in Applied Research and Evaluation Practice? Stewart I. Donaldson, Christina A. Christie and Melvin M. Mark. <http://sites.google.com/site/credibleevidence/> Especially see the free resources, Donaldson, S.I. (2008). A practitioner's guide for gathering credible evidence.

What Constitutes Credible Evidence in Evaluation and Applied Research? Claremont Graduate University Stauffer Symposium

<http://www.cgu.edu/pages/4085.asp>

A Summative Evaluation of RCT Methodology: & An Alternative Approach to Causal Research. Michael Scriven. Journal of MultiDisciplinary Evaluation, Volume 5, Number 9.

<http://jmde.com/>

What is evaluation

One way to address evaluation concerns is by using a collaborative approach to evaluation.

Involve many stakeholders in decisions about the evaluation: how it is going to be conducted, and how the results will be interpreted and used.

Involving stakeholders may:

- reduce suspicion
- increase commitment
- broaden knowledge of evaluation team
- increase the possibility that results will be used.

Additional Resources:

Practical Evaluation of Public Health Programs: Workbook
Public Health Training Network, CDC

<http://www2.cdc.gov/phtn/Pract-Eval/workbook.asp>

What is evaluation

Putting it all together

In sum, planning a program evaluation includes answering these three key points:

1. What, exactly, is the question?
2. How will you get the information to answer the question?
3. What will you do with the answers to the question?

That is, define exactly what you want to find out, plan clearly how to find it out, and have a plan on what to do with the answers.

Additional Resources:

Evaluation: A beginners guide
Amnesty International
<http://www.amnesty.org/en/library/info/POL32/003/1999>

The Program Manager's Guide to Evaluation
Office of Planning, Research and Evaluation
Administration for Children and Families
US Department of Health and Human Services
http://www.acf.hhs.gov/programs/opre/other_resrch/pm_guide_eval

Community Action Resources for Inuit, Métis and First Nations:
Evaluating. Health Canada
http://www.hc-sc.gc.ca/fniah-spnia/pubs/services/_adp-apd/evaluating-evaluation/index-eng.php

Introduction to program evaluation for public health programs: A self-study guide. CDC
<http://www.cdc.gov/eval/framework.htm#formats>

What is evaluation

Copyright by Gene Shackman, July 2008, and February 2010

Some usual legal disclaimers:

This handout can be freely distributed without need for permission, provided it is distributed as is. Distribution for any commercial purpose is strictly forbidden. This handout cannot be sold under any circumstances.

This handout is only for education purposes. It does not represent any guidelines, recommendations or requirements about how to do program evaluation. The only purpose is to provide the general public, consumers, students, and evaluators with information about things that may go into evaluations, so that evaluation may be better understood, and evaluators and clients might work better together to get more out of their evaluation.

In my work on this handout, I do not represent or speak for any organization. I prepared this on my own time, at home, and was not supported by any organization.

I also benefited greatly from feedback from folks on various email lists, and I thank them all!

Materials on web sites listed in this handout do not necessarily reflect my opinions, nor do I assume any responsibility for the content provided at these web sites. This handout only lists web sites with legal content. Listing a website is not necessarily endorsement of any services or organization. The sites are only listed because they have some freely available information. I also do not have any financial relationships with any site or organization listed on this handout.

This handout does not contain or promote; pornography, hatred, racism, propaganda, warez, nudity, hacking activities, violence, degradation, harm or slander. To the best of my knowledge, I do not list any website that contains or promotes any of that either.

The most recent version of this handout is always at
<http://gsociology.icaap.org/methods/>