



# Program Evaluation II: Plans, Designs, and Success

---

Building Capacity for Health Promotion and Chronic Disease Prevention  
6<sup>th</sup> Annual CDC/IUHPE Seminar  
Entebbe, Uganda  
July 8<sup>th</sup> – 18<sup>th</sup> 2009  
Joan Kennelly  
University of Illinois School of Public Health  
[Kennelly@uic.edu](mailto:Kennelly@uic.edu)



# Session Objectives

---

- Review logic model development
- Appreciate the use of logic models in the evaluation process
- Define and explain the purpose of evaluation
- Identify types of evaluations
- Understand evaluation designs and relevant issues of validity and reliability
- Describe key aspects of a successful evaluation





# Logic Models

---

- Illustrate the purpose and content of programs
- Promote development of meaningful evaluation questions
  - Context
  - Implementation
  - Results



# Logic Models

---

Components provide a framework for evaluation questions

Frameworks increase evaluation effectiveness → questions with real value to stakeholders



The Table below describes the relationship between a successful program and the benefits derived from the use of logic models.

<b>Program Elements</b>	<b>Criteria for Program Success<sup>1</sup></b>	<b>Benefits of Program Logic Models<sup>2</sup></b>
Planning and Design	Program goals and objectives, and important side effects are well defined ahead of time.	Finds "gaps" in the theory or logic of a program and work to resolve them.
	Program goals and objectives are both plausible and possible.	Builds a shared understanding of what the program is all about and how the parts work together.
Program Implementation and Management	Relevant, credible, and useful performance data can be obtained.	Focuses attention of management on the most important connections between action and results.
Evaluation, Communication, and Marketing	The intended users of the evaluation results have agreed on how they will use the information.	Provides a way to involve and engage stakeholders in the design, processes, and use of evaluation.

How Logic Models Better Position Programs Toward Success.



# 'Program' Evaluation Questions

---

- What was done?
- To whom?
- How well was it done?
- How much has been done?
- How effective were we?
- What could be done better?
- What should we do differently?

Adapted from CDC



# EVALUATION

---

Attempt to systematically assess the impact of programs or policies on problems they are designed to address →

- data collection and analysis
- dissemination of findings
- action on results



Social program evaluation is a complex and demanding endeavor

Approach with realism, knowledge, and commitment



# Sound Program Evaluation

---

## Required Activities:

1. Analysis of the problem
2. Identification of the goals and outcomes to be evaluated
3. Identification of Standardized Program activities and their measurable objectives
4. Measurement of change
5. Identification of program costs for achieving results

Adapted from Suchman, 1967



---

<b><i>Categories of Change</i></b> (Short, Intermediate and Long Term)	<b><i>Potential Measure</i></b>
Change in knowledge attitude or skill	Individual's behavior
Change in environment	Practices, Programs, Policies
Change in population health	Risk status, morbidity and mortality rates



# Preparing for Evaluation

---

***Scope and purpose of evaluation is provided by the Logic Model***

- Identify the primary questions to be in answered....and in what order
- Highlight the program's theory/logic?
  - Assumptions and 'intellectual landscape' informed by previous research and the literature, as well as the social analysis, epi and environmental assessments



# Preparing for Evaluation

---

- What are the program goals and objectives?
- Who is the primary audience for results?
- How will results be used?

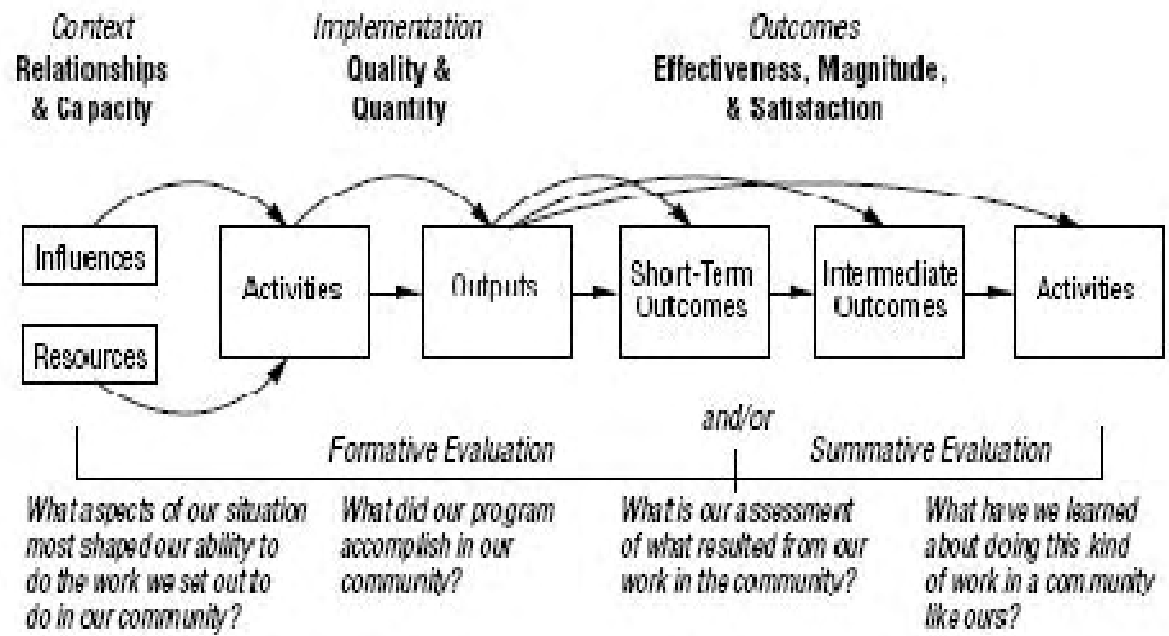


## Types of Evaluation Questions

---

Formative/Process → Improve  
Are we doing things right?

Summative/Outcome → Prove  
Are we doing the right things?





# Evaluation Question Benefits

---

<b>Formative Evaluation – Improve</b>	<b>Summative Evaluation – Prove</b>
Provides information that helps you improve your program. Generates periodic reports. Information can be shared quickly.	Generates information that can be used to demonstrate the results of your program to funders and your community.
Focuses most on program activities, outputs, and short-term outcomes for the purpose of monitoring progress and making mid-course corrections when needed.	Focuses most on program's intermediate-term outcomes and impact. Although data may be collected throughout the program, the purpose is to determine the value and worth of a program based on results.
Helpful in bringing suggestions for improvement to the attention of staff.	Helpful in describing the quality and effectiveness of your program by documenting its impact on participants and the community.

<sup>3</sup> Adapted from Bond, S.L., Boyd, S. E., & Montgomery, D.L.(1997 *Taking Stock: A Practical Guide to Evaluating Your Own Programs*, Chapel Hill, NC: Horizon Research, Inc. Available online at <http://www.horizon-research.com>.



# Evaluation

---

...is more than data collection  
and analysis

...is the groundwork for  
generating knowledge, sharing  
results and moving forward

There is no one best way



# Evaluation Plan

---

- Context: external influences and resources that affect program development and implementation
- Program Assumptions and Conceptual Framework
- Program Goals and Objectives
- Program Outcomes and Expected Impact:
  - Short term
  - Intermediate
  - Long term



# Evaluation Plan

---

## WHO?

### Key Program Informants and Organizational Stakeholders

- Participants and those affected by the program (community residents)
- Program Implementers
- Sponsors, Collaborators, elected officials, advocacy groups



# Evaluation Plan

---

## How?

### Methodological Approach

- Evaluation Design: related to program set up
- Indicator Selection: answer questions related to improving or proving project/program effect
- Data Sources: primary and secondary



# Evaluation Plan

---

How?

Determine Financing

Develop an Implementation Plan

- Roles and responsibilities of key players – internal and external
- Timetable for activities
- Budget



# Evaluation Strategies

---

A program's stage of development affects the evaluation process

Thus, the goals and intent of an evaluation relates to each stage of a program's development, implementation, and completion



<b><i>Stage</i></b>	<b><i>Evaluation Goal</i></b>
Planning – development of new untested activities	Refine as much as possible
Implementation – field testing of activities	Observe/monitor what happens in real world; adjust to improve
Outcome/Effects – adequate time for program effects to emerge	Identify and understand program results including those unintended



# Evaluation Types

---

- Process or Formative
- Outcome or Summative



# Process or Formative Evaluation

---

- Conducted during program implementation
- Describes environmental context of program
- A means of program monitoring and charting progress on meeting objectives
- Generates periodic reports and information to share quickly



# Process or Formative Evaluation

---

- Focus on program activities and outputs
  - What is done by whom
  - Intensity and duration of efforts
  - Details, Details, Details
- Answers question: Was the program implemented as planned? Why? Why not?
- Brings suggestions for improvement to staff attention early



# Process/Formative Evaluation

---

## Benefits:

- Allows for program modification as needed
- Facilitates program replication
- Contributes important information to the outcome evaluation
- Assists in the interpretation of program results



# Outcome or Summative Evaluation

---

- Conducted at either the conclusion of a specific program component period (budget or training); or, at the end of the overall program
- Results based: focus on outcomes (short term and intermediate) and program impact (long term)
  - [ > knowledge/awareness, < use of salt, <use saturated fat, <smoking, >physical activity, <hypertension, >controlled hypertension, ]



# Outcome or Summative Evaluation

---

- Describes quality and effectiveness of program through documentation of impact on participants and the community
  - Is the program more effective than what we usually do?
  - Is the program better than nothing at all?
  - Which of 2 or more programs works better?
  - Is the program efficient?



# Outcome or Summative Evaluation

---

- Intervention EFFICACY: under ideal conditions can the intervention lead to desired outcomes
- Intervention EFFECTIVENESS: when implemented on a wider scale with similar conditions will the intervention lead to the desired outcome
- Intervention COST: developmental and recurrent



# Outcome or Summative Evaluation

---

## Short-term & Intermediate Results

- What effects were observed?
- Can they be attributed to the program?
- Did knowledge, attitude and behaviors change?
- Did the training program achieve its objectives?
- Were there unintended effects?  
Consequences?

Factors that precede and are linked to longer term outcomes.



# Outcome or Summative Evaluation

---

- Long-term Results
  - Relate to overall program goals which are usually more distant and beyond range of the program evaluation....
  - Change in population health status – death rates, incidence or prevalence of particular conditions or risk status
    - Timing related to type of intervention



# Outcome or Summative Evaluation

---

## Identifying Outcomes:

- Realistic and achieved within time-frame of intervention/program
- Sensitive to change and attributable to program

### Examples:

> knowledge, awareness, or skills;

change in behavior - choices or preparation of foods; change in provider practice, or policies; change in social environment - smoking ban, availability of acceptable treated nets;

> availability of foods; > in activity levels; < levels of risk factors



# Outcome or Summative Evaluation

---

- Strategies

- Population or Individually based
- Prospective: need resources committed up front
- Retrospective: limits on measurements and selection of a comparison group



**EVALUATION**

**IS**

**ALL ABOUT COMPARISONS**



# Evaluation Designs

---

Experimental

Quasi-experimental

Non-Experimental

Selection:

- Determined by purpose of evaluation
- Determined by methods used to identify comparison and control groups
- Influenced by the utility of the evaluation information



# Evaluation Designs

---

- Specify what, from whom, and when data are to be collected
  - Data are related to the evaluation questions posed
- Vary in feasibility, cost, validity of results, and degree of selection bias



# Experimental Designs

---

- Random selection and assignment to program and comparison groups
  - Purpose of randomization?
- Traditionally the most desirable; yet, rare to control for all important variables
- Random assignment is also not always feasible or appropriate



# Quasi-Experimental Designs

---

- Follows basic structure of experimental without control for subject differences
- Relatively inexpensive; correlational and less able to impute causation
- Reasons for found associations not always clear
- Comparison groups:
  - constructed by prospective or retrospective matching (identify non-participants with similar characteristics known/believed to influence program outcomes)
  - reflexive comparisons (namely, the change within a person)



# Quasi-Experimental Designs

---

- Cohort (sample from exposure; prospective or retrospective)
- Case-Control (sample from outcome)
- Cross sectional (sample from populations at same point in time, also from different geo-areas)



# Quasi-Experimental Designs

---

- Nonequivalent group posttest only
  - $X \rightarrow O$
  - $\rightarrow O$
- Nonequivalent group pre-test/post-test
  - $O \rightarrow X \rightarrow O$
  - $O \text{----} \rightarrow O$
  - Both of these have advantage of a comparison group
  - If pre-test is similar, group differences are less
- Time series – repeated measures gives trend information  $X \rightarrow O \rightarrow O \rightarrow O$



# Non-Experimental Designs

---

There are no comparison groups

- Single pretest-posttest: compares same group before and after program; cannot distinguish external factors and program effects
- Post-test only
- Single group time-series: looking at change over time; comparison made between 2 points in time. Participants act as their own controls.

Used for descriptive purposes; cannot attribute change to the intervention or claim one intervention better than another..... May be only option



---

Stronger claims of impact success can be made by comparing population outcomes across localities, neighborhoods, or agency units that historically have been comparable with regard to such outcomes and population characteristics



---

Evaluation studies are usually conducted to demonstrate cause and effect

Primary concerns of program evaluation include:

- has the program's theory been translated through clear definitions and measurable indicators
- does design control threats and biases which could undermine the meaningfulness and utility of results → Validity



# Validity

---

- **External Validity** – degree to which evaluation results can be generalized to the population at large
- **Internal Validity** – degree to which successful in creating comparable groups and therefore able to demonstrate a real difference and program effect



# Validity

---

- **Construct Validity** – relates to the logic of measures of the programs social concepts and assumptions
- **Statistical Validity** – basing conclusions on the proper use of statistics



# Validity Varies Along a Continuum

---

## Threats to external validity:

- Population validity → to which population can you apply results; sample characteristics and representativeness
  - Non-random samples such as convenience samples have low external validity
- Ecologic validity → generalizes to realistic settings, beyond a laboratory or tightly controlled community setting
- Multiple treatment effects → participants involved in multiple programs
- Situation effects → influence of resources provided for the 'demonstration' that are needed for wider implementation



# Validity Varies Along a Continuum

---

- Internal validity defends against sources of bias that would effect the proposed cause-effect processes being studied
- < internal validity means that other factors may be responsible for the observed 'program' effects

Threats to internal validity?



# Validity Varies Along a Continuum

---

- Threats to internal validity:  
Selection bias, Mortality bias, History  
Maturation, Instrumentation and Testing  
Regression to the mean



- 
- Sound evaluation designs control (to the extent feasible) threats to validity
  - Best to have a control/comparison group



---

Evaluation designs specify  
what, from whom, and when  
data are to be collected



# Evaluation Data: Indicators

---

Quantified measurements repeated over time to track achievement of objectives

- Valid - accurate measure of what is being evaluated
- Reliable – measures consistently and dependably over time
- Sensitive – will pick up change due to program
- Utility – data easily interpreted and guides project change



# Evaluation Data: Indicators

---

- The selection of indicators is a good starting point for designing data collection and reporting strategies
- Well defined indicators make it easy to collect high quality data
- Several indicators are usually needed to track implementation and effects of a complex program or intervention



## Evaluation Data: Sources of Indicators/Evidence

---

- People
- Documents
- Surveillance data
- Surveys
- Vital Records
- Observation

More than one source is recommended to provide a comprehensive review



# Evaluation Data: Reliability and Validity of Measurement

---

Reliability – extent to which there are consistent results on independent measurements

- Types of reliability:

- Test-retest → same observer, subjects, instrument, conditions, different times
- Inter-rater → different observer, same subjects, instrument, conditions and times



# Evaluation Data: Reliability and Validity of Measurement

---

Validity – extent to which factor of interest is successfully captured

Example:

salt use – self report, urine samples, shaker weight, sales

physical activity – self report, pedometers, accelerometers, fitness tests, mets

smoking – self report, sales, cotinine, blood nicotine



---

## Data analysis should match your program design

- Comparing groups, comparing points in time, comparing to a standard, assessing changes in knowledge attitudes and behavior.....
  - Mean differences, absolute differences, % change, rate of change, trends...and when appropriate test for statistical significance



## Evaluation of Social Programs: Bottom Line

---

- Problems are complex and multilevel
- No single study will ever be sufficient for understanding a phenomenon with complete validity
- Explication of underlying theory is critical
- Need to attend to process and outcomes



## Evaluation of Social Programs: Bottom Line

---

- Multiple perspectives are crucial
- Triangulation of available information important – using multiple measures for a specific domain or dimension.
- Multiple and varied sources of evidence within a particular study are also needed: agreement across stakeholders on what is success - <cost, <risk factors, > satisfaction
- Strive for comparable groups relatively free of bias



# Ensuring Successful Evaluations

---

- Build evaluation concept into the planning of the program
- Establish routine information system
- Include evaluation activities in the project budget
- Hold regular monitoring and evaluation meetings with project staff
- Encourage review and revision of evaluation plan



# Ensuring Successful Evaluations

---

- Design and methods are appropriate for questions being asked
  - Project is planned on basis of thorough review of evidence from epidemiologic, behavioral, and social research and problem analysis/community assessment
- Reasonable control of bias
- Adequate sample size
- Intervention (quality and duration) has an 'effect' detected above background noise of community changes.
- Measurement is reliable and valid



# Ensuring Successful Evaluations

---

Note:

If you plan to include in the evaluation, a focus or reporting of personal information of participants in the program/intervention, you may need their informed consent

Check with you local institutions.



# References

---

- Logic Model Development Guide Evaluation Handbook, W.K.Kellog Foundation  
<http://www.wkkf.orh>
- Eric Clearinghouse on Assessment and Evaluation  
<http://www.ericae.net>
- United Way of America, Measuring Program Outcomes: A Practical Approach  
<http://www.unitedway.org/outcomes/contents.htm>
- <http://www.CDC.gov>
- Health Program Planning: An Educational and Ecological Approach 4<sup>th</sup> edition, L.W.Green and M.W.Kreuter. McGraw Hill, New York 2005