Showing that your work matters:
Usable program evaluation for the nonprofit sector

Rob Fischer, Ph.D.
Center on Urban Poverty and Community Development

Raymond John Wean Foundation Summit ’15
September 24, 2016
The Iron Triangle of Program Evaluation

Pick 2!
Fast + Cheap = Narrow
Cheap + Broad = Slow
Fast + Broad = Expensive
Faced with evaluation demands...
Evaluation purposes

• Program planning and improvement
  → formative evaluation

• Accountability
  → summative evaluation

• Knowledge generation
3 Key Roles for Evaluation

• Clarifying the program model
• Evaluating program process
• Evaluating program outcomes
Clarifying the Program Model

- Understanding the underlying need
- Defining the target population
- Designing the intervention/response
- Selecting the point of intervention
- Naming the “active ingredients”
## Getting to ‘results’

### Domains of Results

<table>
<thead>
<tr>
<th>Goal</th>
<th>Process results</th>
<th>Output results</th>
<th>Outcome results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convert resources to</td>
<td>Program delivery reaches targets</td>
<td>Targets exhibit desired changes</td>
<td></td>
</tr>
<tr>
<td>program delivery</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Example Indicators       |                                         |                                     |                                         |
| Staff hired              |                                         |                                     |                                         |
| MOU signed               |                                         |                                     |                                         |
| Materials developed      |                                         |                                     |                                         |

| Indicator                |                                         |                                     |                                         |
| Staff hired              |                                         |                                     |                                         |
| MOU signed               |                                         |                                     |                                         |
| Materials developed      |                                         |                                     |                                         |

| Evaluation approach      |                                         |                                     |                                         |
| Process evaluation       |                                         |                                     |                                         |

| Evaluation approach      |                                         |                                     |                                         |
| Outcome evaluation       |                                         |                                     |                                         |

### Example Indicators

- Staff hired
- MOU signed
- Materials developed

### Evaluation approach

- Process evaluation
- Outcome evaluation
The basics of logic models

• Logic model should simplify the who, what, where, when, how, and why of the program
• Identify the inputs, activities, outputs, and outcomes
• Remember: outcomes must be changes in the clients served by the program
• Program theory = rationale for intervening in the way the program does (with whom and when)
Exhibit A

Summary of Program Outcome Model

INPUTS → ACTIVITIES → OUTPUTS → OUTCOMES

**Resources dedicated to or consumed by the program**
- e.g.: money, staff and staff time, volunteers and volunteer time, facilities, equipment and supplies

**What the program does with the inputs to fulfill its mission**
- e.g.: feed and shelter homeless families, provide job training, educate the public about signs of child abuse, counsel pregnant women, create mentoring relationships for youth

**The direct products of program activities**
- e.g.: number of classes taught, number of counseling sessions conducted, number of educational materials distributed, number of hours of service delivered, number of participants served

**Benefits for participants during and after program activities**
- e.g.: new knowledge, increased skills, changed attitudes or values, modified behavior, improved condition, altered status

Program: **Family Development Center**

**Inputs**
Agency provides 1 center director, 1 social worker, 1 resident manager couple, and 1 assistant resident manager; 14 unit efficiency apartment facility; financial resources for day care, job training, etc.

**Activities**
Transitional housing program for homeless women with their first child (under age 1)

**Outputs**
- # families in residence; # group sessions; # counseling sessions; # referrals

**Initial Outcomes**
- Women obtain employment
- Families gain access to needed public assistance
- Families exhibit decreased reliance on public assistance
- Women gain knowledge of parenting, homemaking and coping skills
- Women avoid second pregnancy

**Intermediate Outcomes**
- Women use effective parenting, homemaking and coping behaviors
- Women continue to use appropriate family planning practices

**Longer-term Outcome**
Families continue progress toward economic self-sufficiency
PARENT EDUCATION PROGRAM

Situation:
During a county needs assessment, a majority of parents reported they were having difficulty parenting and suffering stress as a result.

INPUTS

- Staff
- Money
- Partners
- Research

OUTPUTS

Assess parent ed programs
Design-deliver evidence-based program of 8 sessions
Facilitate support groups

Parents of 3-10 year olds attend

OUTCOMES

Parents increase knowledge of child dev
Parents better understanding their own parenting style
Parents gain skills in new ways to parent
Parents gain confidence in their abilities
Parents identify appropriate actions to take
Parents use effective parenting practices

Reduced stress
Improved child-parent relations
Strong families

DEVELOPING A LOGIC MODEL: TEACHING AND TRAINING GUIDE  2/29/2008
© 2008 by the Board of Regents of the University of Wisconsin System. All rights reserved.

JACK, JOSEPH AND MORTON MANDEL
SCHOOL OF APPLIED SOCIAL SCIENCES
CASE WESTERN RESERVE UNIVERSITY
Logic model essentials

• A logic model is a thumbnail sketch of a program, and cannot exist without a supporting narrative

• If there is not agreement on what the program is at this level, this suggests bigger issues

• Distinctions
  – Proximal vs distal outcomes
  – Outcomes vs effect vs impact
Thinking about outcomes over time

**Shorter-Term Outcomes**
- Achieved during program timeframe
- Within program control
- “expect to see”

**Intermediate Outcomes**
- Achieved at the end / beyond program timeframe
- Follow shorter-term outcomes
- “want to see”

**Longer-Term Outcomes**
- Achieved after program timeframe
- Outside direct program control
- “hope to see”

Fig. 1: Outcomes Chain
Program: Home Alone Home Visiting Program

Long-term Outcomes
- Women and their children avoid future criminal activity

Intermediate Outcomes
- Women complete the program
- Women gain knowledge about healthy parenting behaviors

Initial Outcomes
- Women participate in the program
- Women practice healthy parenting behaviors after participating in program

Outputs
- Funder is satisfied with the way the money is being spent and media coverage of program is positive

Activities
- Home visiting program that provides nutrition information and parenting training to expectant young women

Inputs
- Agency provides 6 staff; 2 office locations; 200 clients participate in the program
Monitoring program outcomes

• Guidelines for outcome indicators
  – Focus on changes in program participants
  – Use measures of pre to post change
  – Select outcomes closely linked to program
  – Always consider client self-report

• Pitfalls in outcome monitoring
  – Funder’s key outcomes
  – Corruptibility of indicators
  – Interpret indicators in programmatic and social context
Break-out # 1

• Discuss a selected program example and work to flesh out a preliminary logic model

• Focus on (you can skip inputs and activities)
  – Outputs
  – Outcomes – initial and beyond

• At this stage identify the right outcomes for the program whether or not you will be able to measure them

• Use arrows to show conceptual linkages between outcomes
Evaluating Program Process

- Measuring program efforts
- Articulating a sense of dose
  - Sequencing
  - Duration
  - Intensity

Figure 5.2 Quality enhancement visits received by providers within 12 months of certification
Evaluating Program Outcomes

• Single Group (the served)
  – Outcome monitoring – KAB + S
  – Longitudinal (e.g., time series)

• Comparative
  – Nonequivalent comparison groups
  – Matched designed
  – Unit-based designs (>individual)
Choosing evaluation methods

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Existing Records</th>
<th>Focus Groups</th>
<th>Interviews</th>
<th>Observation</th>
<th>Portfolio/Journal</th>
<th>Tests</th>
<th>Written surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anonymous</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Cost-effective</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Easily administered</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Generalizable</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Quantitative</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Probing (allows)</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rich data (collects)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>
Differences in perspective

Level of Behavior Problems Rated by Parents (P) & Teachers (T) (n=105)

Boys-T
- Pre: 12.05
- Post: 13.30

Boys-P
- Pre: 12.74
- Post: 13.58

Girls-T
- Pre: 8.08
- Post: 8.62

Girls-P
- Pre: 10.06
- Post: 12.32

Score on Behavior Problems Index
Making use of performance data

- Program completed
- 30 day retention
- 90 day retention
- 180 day retention

% of cases

Agencies: Agency 1, Agency 2, Agency 3, Agency 4, Agency 5, Agency 6, Agency 7, Agency 8, Agency 9, Agency 10
Ex: Exploring data sources

• An RFP has been issued for developing a program monitoring approach for a case management for individuals who are homeless and have substance abuse problems.

• What are the principal strengths and limitations of the following data sources in monitoring program implementation in this case?
  – data collected by the evaluator
  – service utilization records (e.g., case records, MIS)
  – program participants
Developing measures

• Valid measures
  – Measures the behaviors that program is designed to change

• Reliable measures
  – Characteristics of observation procedures
  – Different observer/same observers reporting similar levels of a variable

• Sensitivity to change
  – Does the measure capture the change?

• Cost-effective measures
Data collection

• Data sources
  – Clients, existing records (internal or external), community, trained observers, mechanical measures

• Instrumentation
  – Survey
  – Interview or focus group
  – Data extraction
Surveys

• Mode of administration
  – Paper & pencil (in-person or mail)
  – Phone

• Origin
  – Standardized vs. “Home-made”
  – Trade-off: cost/ease of use -- validity/reliability
Extraction from records

• Extracted by staff
  – Closed case or termination summary
• Extracted by evaluator
  – Record review
• Electronic data extract

In both cases, use standardized coding sheet to extract elements
Minimum data collection elements

- Client background-related - demographics
- Service-related - Amount & type of service received
- Outcome-related - Changes in attitude, knowledge, or behavior
- Cover letter/introduction
- Question types
  - Fixed-choice - Scaling options: dichotomous, multi-category, ordinal; Anchored scales
  - Fill-in and Open-ended
Common problems

• Readability
  – Reading level too high
  – Use of jargon, vague words

• Response issues
  – Double-barreled questions
  – Overlapping response categories

• Lack of objectivity
  – Leading questions
Other issues

• Use of informed consent
• Payment for participation
• Anonymity and confidentiality
• Response rate
  – What to expect and dealing with low response
  – Methods for improving
Break-out # 2

• Return to your logic model and focus on the task of data collection

• For each output and outcome, identify
  – Data source
  – Data method
  – Timing of data collection

• Are there opportunities to compare these to data on unserved populations?
Engaging clients in evaluation

• We send a message about the importance of client voice in both the program model and the evaluation model
  – Clients as the program target
  – Clients as the program beneficiary
  – Clients as one of a number of stakeholders for the evaluation
  – Clients as a data source for the evaluation
### Tiers of Client Involvement in Evaluation

<table>
<thead>
<tr>
<th>Tiers</th>
<th>Client Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governing</td>
<td>Clients provide leadership on evaluation approach/direction</td>
</tr>
<tr>
<td>Advisory</td>
<td>Clients provide input/feedback on evaluation approach/direction</td>
</tr>
<tr>
<td>Direct/Incidental</td>
<td>Clients provide data that are used in the evaluation approach itself</td>
</tr>
<tr>
<td>Proxy</td>
<td>Client views represented by others in evaluation (e.g., program staff)</td>
</tr>
<tr>
<td>Absent</td>
<td>Client views not meaningfully represented in evaluation</td>
</tr>
</tbody>
</table>
What are the limitations of outcome measurement?

• Cannot determine causation (program $\rightarrow$ effect)
• Making comparisons between programs and/or within a program over time can be problematic
• Outcome data do not tell us
  – why outcome changes do or do not occur
  – how to improve the program
How much of the effect is due to the program?

Desired Outcome ($Y$)

Pre | Time ($X$) | Post

T

C

Net Effect

Gross Effect
Types of designs

One group posttest only design

\[ T_2 = \text{Posttest} \]
\[ P = \text{Program or intervention} \]
Types of designs

Before and After Design

One group pretest-post-test design

$T_1$ = Pretest (treatment group)
$T_2$ = Posttest (treatment group)
$P$ = Program or intervention
2. Comparison Group Design

- \( T_1 \) = Pretest (treatment group)
- \( T_2 \) = Posttest (treatment group)
- \( P \) = Program or intervention
- \( C_1 \) = Pretest (comparison group)
- \( C_2 \) = Posttest (comparison group)
Compared to what?

✓ Over time
  ✓ Pre to post
  ✓ Longitudinal

✓ Between groups
  ✓ Randomly composed
  ✓ Naturally occurring
    (waitlist, other programs)
  ✓ National norms/standards

Low Ability to Attribute Effect

Post-test only  Pre & Post test  Nonequivalent comparison group  Quasi-experiment

High Ability to Attribute Effect

Randomized experiment
Recommendations

• Be vigilant for opportunities to use more rigorous comparative designs
• Add data points over time -> show outcome trends
• Add data sources/methods -> triangulate
• Engage stakeholders in program design and evaluation -> increase authenticity
• Disseminate findings, including via publication
Internet Resources

Kellogg Logic Model Development Guide:  
http://www.wkkf.org/Pubs/Tools/Evaluation/Pub3669.pdf

Logic Model Workbook:  
http://www.innonet.org/?section_id=62&content_id=143

United Way - Outcome Measurement Resources:  
http://national.unitedway.org/outcomes/library/pgmomres.cfm

Logic Model Course:  
www.uwex.edu/ces/lmcourse

Logic Model Templates:  
http://www.uwex.edu/ces/pdande/evaluation/evallogicmodelworksheets.html