



# An Empirical Review of Theory- Driven Evaluation Practice from 1990-2008

Chris L. S. Coryn, Lindsay A. Noakes,  
Daniela C. Schröter, & Carl D. Westine

The Evaluation Center, Western Michigan University

# Research on Evaluation

---

- Largely led by Christina Christie, J. Bradley Cousins, and the work of a few others
- Some recent examples
  - Empowerment (Miller & Campbell, 2006)
  - Participation (e.g., Cousins & Whitmore, 1998; Cullen, 2009)
  - Standards (Wingate, 2009)
  - Use (e.g., Brandon & Singh, 2009; Cousins & Leithwood, 1986; Johnson, Greenesid, Toal, King, Lawrenz, & Volkov, 2009; Shulha & Cousins, 1997)
- This study sought to (1) identify the central tenets/core principles of theory-driven evaluation and (2) assess the degree to which published case examples are congruent with these tenets/principles

# Core Principles

---

- Five central tenants/principles and 17 subprinciples (see pp. 11-12 for greater detail)
  1. Formulate a plausible program theory
  2. Formulate and prioritize evaluation questions
  3. Use program theory to guide design, planning, and conduct of the evaluation
  4. Measure constructs (process, outcome, context) postulated by program theory
  5. Identify breakdowns, side effects, determine program effectiveness (or efficacy), and describe and explain cause-and-effect associations between theoretical constructs

# Guiding Research Questions

---

1. In what kinds of settings, with what populations, of what scale and scope, and for what purposes are theory-driven evaluations conducted?
2. Why do evaluators and/or their collaborative partners choose theory-driven evaluation as their evaluation strategy?
3. How are stakeholders involved in theory-driven evaluations and in what phases of the evaluation?
4. To what degree are theory-driven evaluations tailored (e.g., investigating specific aspects of a program theory, answering particular questions) versus comprehensive?
5. To what extent are the core principles of theory-driven evaluation evident in theory-driven evaluation practice? What types of evidence for supporting cause-effect inferences are used in theory-driven evaluation practice?

# Method

---

- Sample
  - Multistage sampling design
    - 181 total articles, books, and book chapters identified (over 3 sampling stages)
    - Final sample of  $N = 39$  (21% of 181 identified) “codable” case examples
- Data analysis
  - Two stages
    - Each chapter and article randomly assigned to six groups of two coder-pairs who worked independently (coefficient of agreement = .83 across all coded units)
    - Consensus of coders (used for data analysis)

# Results: Setting and Populations

---

- Settings
  - Health = 46.2%
  - Education = 25.6%
  - Crime and safety = 7.7%
  - Transportation = 7.7%
  - International development = 5.1%
  - Environmental affairs = 5.1%
  - Business = 2.6%
- Target populations of programs evaluated
  - Children = 41.1%
  - College students = 12.8%
  - Adolescents and young adults, general populations, low-income, and “other” = 10.3% (each)
  - Working adults = 5.1%

# Results: Scale/Scope and Purpose

---

- Scale and scope
  - Small local = 33.3%
  - Large local = 23.1%
  - Small regional = 2.6%
  - Large regional = 10.3%
  - Small national = 10.3%
  - Large national = 12.8%
  - Small international = 2.6%
  - Large international = 5.1%
- Purpose
  - Summative only = 48.7%
  - Both formative and summative = 38.5%
  - Formative only = 10.3%
  - Unclear = 2.6%

# Results: Theory Formulation

---

- Theory formulation process
  - Predominately existing scientific theory (87.2%) and/or theories and assumptions held by stakeholders (48.7%, n = 19)
  - Program observation = 15.4%
- Comprehensive versus tailored
  - 59.0% investigated the postulated program theory comprehensively
  - 41.0% investigated only one specific aspect of the postulated program theory such as the program's process theory, outcome or impact theory, or one particular causal chain



# Results: Stakeholders Engaged

---

- Stakeholder groups engaged
  - Policy makers at 15.4%
  - Funders = 20.5%
  - Program staff = 35.9%
  - Direct program impactees = 17.9%
  - Indirect program impactees (i.e., those affected by, but not directly receiving program services) = 5.1%
- Average number of stakeholder groups engaged = 1

# Results: Phases of Engagement

---

- Phases in which stakeholders are engaged
  - Not explicitly described = 56.4%
  - **Initial theory formulation = 38.5%**
  - Question formulation/prioritization = 12.8%
  - Evaluation design = 7.7%
  - Data collection = 7.7%
  - Data analysis = 2.6%
  - Interpretation of results = 7.7%
  - Dissemination of evaluation results = 2.6%
- Average number of phases in which stakeholders are engaged = 1

# Results: Practice-Theory Match

- Degree to which subprinciples within core principles #2, #3, and #4 (use of program theory to...) were met

Number of Subprinciples Met	Number of Cases	Percent of Cases	Formulate Questions	Prioritize Questions	Determine Evaluation Design	Measure Process Constructs	Measure Outcome Constructs	Measure Contextual Constructs
0	1	2.6%	0	0	0	0	0	0
1	4	10.3%	0	0	1	1	2	0
2	8	20.5%	6	0	0	2	6	2
3	8	20.5%	6	0	2	6	7	3
4	11	28.2%	10	1	7	11	11	4
5	5	12.8%	5	3	5	5	4	3
<b>6</b>	<b>2</b>	<b>5.1%</b>	2	2	2	2	2	2
<hr style="border-top: 1px dashed black;"/>								
Total	39	100.0%	29	6	17	27	32	14
Percent of Total			74.4%	15.4%	43.6%	69.2%	82.1%	35.9%

# Results: Practice-Theory Match

- Degree to which subprinciples within core principle #5 were met

Number of Subprinciples Met	Number of Cases	Percent of Cases	Identify Breakdowns	Identify Side Effects	Causal Description	Causal Explanation	Moderating Variables	Mediating Variables
0	3	7.7%	0	0	0	0	0	0
1	4	10.3%	1	0	3	0	0	0
2	1	2.6%	1	0	1	0	0	0
3	5	12.8%	0	0	5	5	1	4
4	9	23.1%	3	0	9	9	6	9
5	13	33.3%	13	2	13	13	11	13
<b>6</b>	<b>4</b>	<b>10.3%</b>	4	4	4	4	4	4
<hr style="border-top: 1px dashed black;"/>								
Total	39	100.0%	22	6	35	31	22	30
Percent of Total			56.4%	15.4%	89.7%	79.5%	56.4%	76.9%

# Results: Practice-Theory Match

- Total subprinciples met

Combined Subprinciples Met	Number of Cases	Percent of Cases	
0	0	0.0%	•10 cases (25.6%) applied six or fewer (less than half)
1	1	2.6%	
2	1	2.6%	
3	1	2.6%	
4	3	7.7%	
5	2	5.1%	
6	2	5.1%	
7	5	12.8%	
8	4	10.3%	
9	8	20.5%	
10	4	10.3%	•12 cases (30.8%) applied 10 or more
11	5	12.8%	
12	3	7.7%	
<b>13</b>	<b>0</b>	<b>0.0%</b>	
Total	39	100.0%	

# Implications for Theory/Practice

---

- Need for published examples (in particular, in international development)
- Greater detail describing the process of theory development, how stakeholders are engaged, and how questions are formulated and prioritized and by whom, etc.
- Clear identification of motives for selecting the approach over other approaches, even if this rationale is ideological preference
- Participation of more than one stakeholder group in more than one area or phase of the evaluation
- More attention needs to be given to examining side effects and unintended outcomes by theory-driven evaluators