

Evaluation Theory: The Enduring Issues

Miles McNall, PhD

Assistant Director

Community Evaluation and Research Center

University Outreach and Engagement

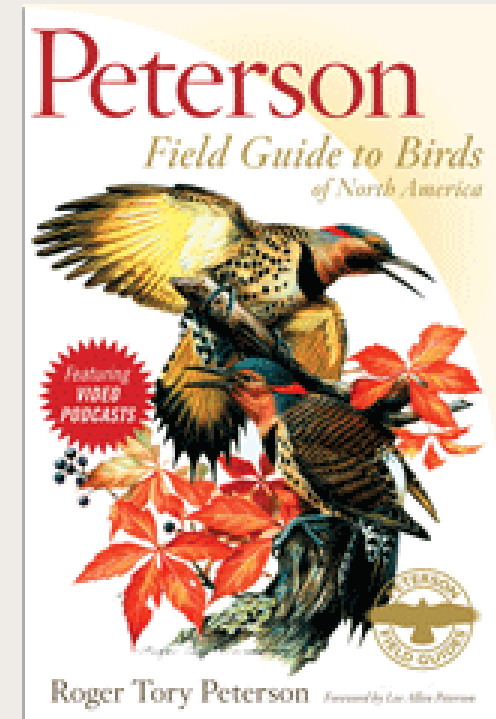
Evaluation Circle

October 30, 2009



A Field Guide to Evaluation Theory

- Typical evaluation field guides: typologies of different models for evaluation practice
- This field guide: the issues that evaluators are likely to encounter in their evaluation practice.



Overview of Workshop

- ❖ Introductions
 - Evaluation and program evaluation
 - Evaluation theory
- ❖ Evaluation theories that have influenced you
 - Typologies of evaluation approaches
 - Enduring issues in evaluation
- Break
 - Enduring issues in evaluation
- ❖ Enduring issues in your evaluation practice



Evaluation, Program Evaluation, & Theories of Evaluation Practice



E-*valu*-ation

- Determines the “merit, worth, or significance” of something. (Scriven, 1967)
 - Includes the evaluation of products, personnel, programs, etc...



Program Evaluation

“The use of social research procedures to systematically investigate the effectiveness of social intervention programs...” (Rossi, Freeman and Lipsey, 1999, p.2)



Objections to the R, F & L Definition

- *All* program effects should be assessed, intended or not
- *How* programs operate is just as important
- Stakeholder questions are paramount
- There are no objective program effects



Evaluation Theory

- Evaluation “theory” is not really theory.
- A theory is a body of principles offered to explain a phenomenon
- More accurate term: scholarship on approaches to evaluation practice
- Less cumbersome term: *theories of evaluation practice*



Theories of Evaluation Practice

- Purpose: *“to specify feasible practices that evaluators can use...”* (Shadish, Cook and Leviton, 1991, p. 36).



Influential Evaluation Approaches

- With what evaluation approaches or theorists are you most familiar?
- What issues in evaluation seem central to these approaches or theorists?



Stufflebeam's Evaluation Models

Evaluation Models:

- *“...approaches that assist evaluators in designing and carrying out useful, defensible program evaluations” (Stufflebeam & Shinkfield, 2007, p. 59).*



Stufflebeam's Evaluation Models

Three major types of models:

- Questions and/or methods-oriented
- Improvement/accountability-oriented
- Social agenda/advocacy approaches



Questions/Methods-Oriented Approaches

- *Quasi-evaluation studies*
 - narrowly focused
 - do not to assess overall merit and worth
- Two sub-types:
 - *Questions-oriented*: Begin with a narrowly-defined set of questions
 - Example: Objectives-based studies
 - *Methods-oriented*: Use a particular method or combination of methods
 - Example: Experimental studies



Improvement/Accountability-Oriented

- “True” evaluation studies that assess merit and worth in a comprehensive, rigorous manner
- Three major sub-types
 - Decision/accountability-oriented
 - Stufflebeam’s CIPP (Context, Input, Process, Product) Model
 - Consumer-oriented
 - Scriven’s Goal-Free Evaluation Model
 - Accreditation/certification
 - JCAHO accreditation studies



Social Agenda/Advocacy Approaches

- Aimed at improving lives of disadvantaged or marginalized groups
- Draw on perspectives of stakeholders in characterizing, investigating, and judging programs
- Favor a constructivist orientation and use of qualitative methods
- Provide for democratic engagement of stakeholders in obtaining and interpreting findings



Social Agenda/Advocacy: Major Types

- Client-centered studies (Stake)
- Constructivist evaluation (Guba and Lincoln)
- Deliberative democratic evaluation (House)
- Utilization-focused evaluation (Patton)



The Evaluation Theory Tree

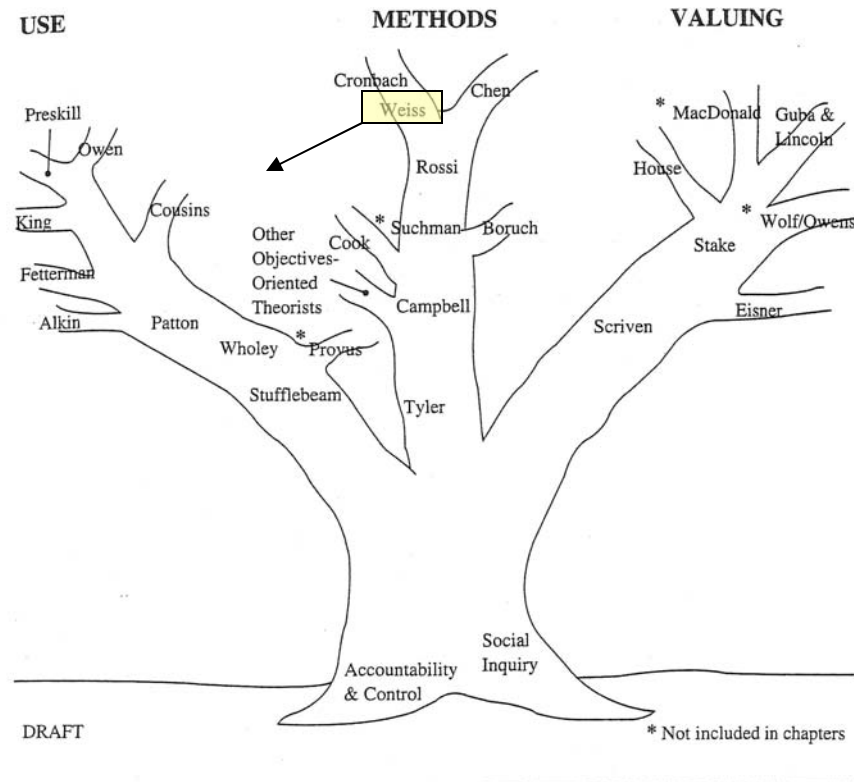


Figure 2.1 Evaluation Theory Tree

Alkin, M. 2004



Limitations of Evaluation Typologies

- Divide evaluation field into opposing camps
- Obscure issues that are common across all evaluations
- Hinder the development of a flexible and responsive approach to evaluation practice
- Give short shrift to certain evaluation issues



Theories of Evaluation Practice

Purpose:

*“to specify feasible **practices** that evaluators can use to **construct knowledge** of the **value of social programs** that can be **used** to ameliorate the social problems to which programs are relevant” (Shadish, Cook and Leviton, 1991, p. 36) [emphasis mine].*



Enduring Issues in Evaluation Theory

Social Programming

How do programs develop and change?
How are programs connected to social problem solving?

Knowledge Construction

How can research construct valid or credible knowledge about programs?

Valuing

How can justifiable value statements about programs be constructed?

Use

How can findings can be used to modify programs?

Practice

What are effective strategies and tactics for conducting evaluations?



An Additional Issue in Evaluation Theory

*Purpose or aim of
program evaluation*

What is are the purposes of
conducting program
evaluations?

Other issues in evaluation?



Enduring Issues in Evaluation Theory

➤ *Purpose of Evaluation*

Social Programming

Knowledge Construction

Valuing

Use

Practice



Evaluation Purpose

- What is evaluation's role in the larger context of policy making and social problem solving?
 - Views of a social problem
 - Knowledge of past attempts to ameliorate it
 - Policy development
 - Program design
 - Program implementation
 - Program modification
 - Program dissemination or termination



A Lesson in Humility

Evaluation is fallible

Evaluation is but one source of evidence

Evidence is but one input into policy

Policy is but one influence on practice

Practice is but one influence on outcomes

-Weiss, Murphy-Graham, & Birkeland, 2005



Trends in Evaluation Purpose

Summary
Judgment



Multiple
Purposes

Evaluation renders
summary
judgments about
the effectiveness of
programs

Evaluation
serves several
purposes



Purposes of Program Evaluation: The Golden Oldies

- Two types of evaluation (Scriven, 1967):
 - *Formative*: Conducted for the purpose of improving a program
 - *Summative*: Conducted for the purpose of rendering a judgment on program merit, worth, or significance
- *When the cook tastes the soup, that's formative.*
When the guests taste it, that's summative!
 - Robert Stake



Four Purposes of Program Evaluation

- *Assessment of merit and worth.* (Summative).
- *Program and organizational improvement.* (Formative).
- *Oversight and compliance.* Assessment of the extent to which programs comply with statutes, regulations, or other mandates.
- *Knowledge development.* Development and testing of general theories about social processes as they occur in the context of social policies and programs.



(Mark, Henry, & Julnes, 1999)

Social Problem Solving and Evaluation Purposes

<i>Social Problem Solving</i>	<i>Evaluation Purpose</i>
<ul style="list-style-type: none">• Views of a social problem	<ul style="list-style-type: none">• Knowledge development
<ul style="list-style-type: none">• Knowledge of past attempts to ameliorate social problem	<ul style="list-style-type: none">• Assessment of merit and worth
<ul style="list-style-type: none">• Policy formation	<ul style="list-style-type: none">• Knowledge development• Assessment of merit and worth
<ul style="list-style-type: none">• Program design	<ul style="list-style-type: none">• Knowledge development• Assessment of merit and worth
<ul style="list-style-type: none">• Program Implementation	<ul style="list-style-type: none">• Oversight and compliance
<ul style="list-style-type: none">• Program modification	<ul style="list-style-type: none">• Program and organizational improvement
<ul style="list-style-type: none">• Program replication or termination	<ul style="list-style-type: none">• Assessment of merit and worth



Enduring Issues in Evaluation Theory

Purpose of Evaluation

➤ *Social Programming*

Knowledge Construction

Valuing

Use

Practice



Social Programming Theory

- Key Issues
 - How do programs develop and change?
 - How are programs connected to social problem solving?
- Assumption
 - Social problems can be ameliorated through better social programming
 - Contrast with argument of *Losing Ground*



Social Programming Theory

- Good social programming theory considers:
 - How social policies and social programs change
 - How program change contributes to social problem solving
 - How external constraints shape programs
 - Internal program structure and functioning



Trends in Theories of Social Programming

The
Experimenting
Society



The Tinkering
Society

Testing out bold,
innovative
approaches to
social problem
solving

Making marginal
changes to existing
program



Social Programming Theory: Stage One

The Historical Context:

- Lyndon Johnson's Great Society social programs
 - Civil rights, War on Poverty, Medicare and Medicaid
- Optimism about:
 - Potential of social programs to ameliorate stubborn social problems
 - Capacity of social science to develop effective social interventions and assess their effectiveness
 - Adoption of demonstrably effective programs and the termination of evident failures by policy makers



Donald Campbell's Experimenting Society

- In which we:
 - *try out new programs designed to cure specific social problems, in which we learn whether or not these programs are effective, and in which we retain, imitate, modify or discard them on the basis of their apparent effectiveness (Campbell, 1969, p. 409).*



Disillusionment: Disappointing Results

- *“Soon results began to pour in. The news was dismaying. Nothing seemed to be working as expected” (Weiss, 1987, p. 41).*
- *Iron Law of Evaluation: “The expected value of any net impact assessment of any social program is zero” (Rossi, 1985, p.2)*



Disillusionment: Methodological challenges

- Absence of control of confounding variables,
- Un-measurable program goals,
- Lack of program staff cooperation,
- Resistance to random assignment,
- Lack of standardization of treatment across projects, individuals, times



Disillusionment: The Political Context

- Programs are subject to political pressures that can serve to protect ineffective programs
- Evaluation findings are one of many inputs into decisions about programs
- Evaluation is inherently conservative



A More Realistic View

- Increasingly incremental approach to social programming
- Evaluation of marginal changes/additions to existing programs
- More sophisticated view of
 - realities of social programming
 - political context of evaluation
 - limits of research methods



Enduring Issues in Evaluation Theory

Purpose of Evaluation

Social Programming

➤ *Knowledge Construction*

Valuing

Use

Practice



Knowledge Construction Theory

- Key issue
 - How can credible knowledge of programs be created?
- Key assumptions
 - the nature of reality (ontology)
 - the bases of claims to knowledge (epistemology)
 - how knowledge is constructed (methodology)



From Post-Positivism to Multiple Paradigms

Dominance of Post-positivism

Post-positivism:

human knowledge is *conjectural*, but these conjectures have warrants (justifications), which can be modified or withdrawn in the light of further investigation.



Respect for multiple paradigms (?)

A respectful dialogue among paradigms

- Post-positivism
- Constructivism, post-modernism, etc.



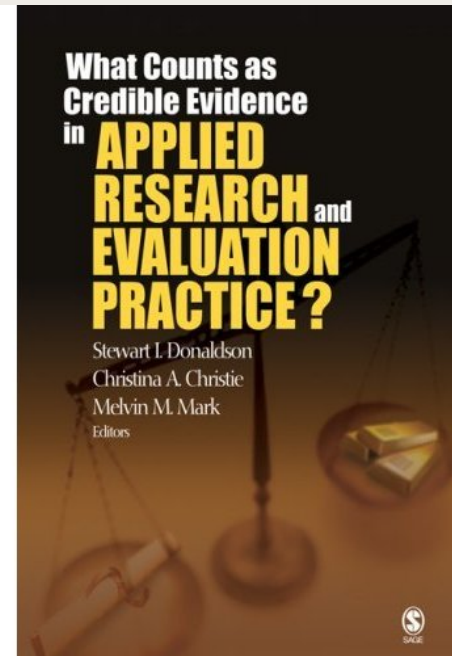
Paradigm Wars in Evaluation

- AEA statement on prioritization of RCTs in educational research
 - RCTs are not the only methods capable of generating understanding about causality
 - RCTs can sometimes produce misleading results
 - RCTs can be viewed as unethical
 - Data sources are often inadequate for RCTs



What Counts as Credible Evidence

- The AEA policy statement controversy in the words of the participants
- A more general discussion of what constitutes credible evidence.



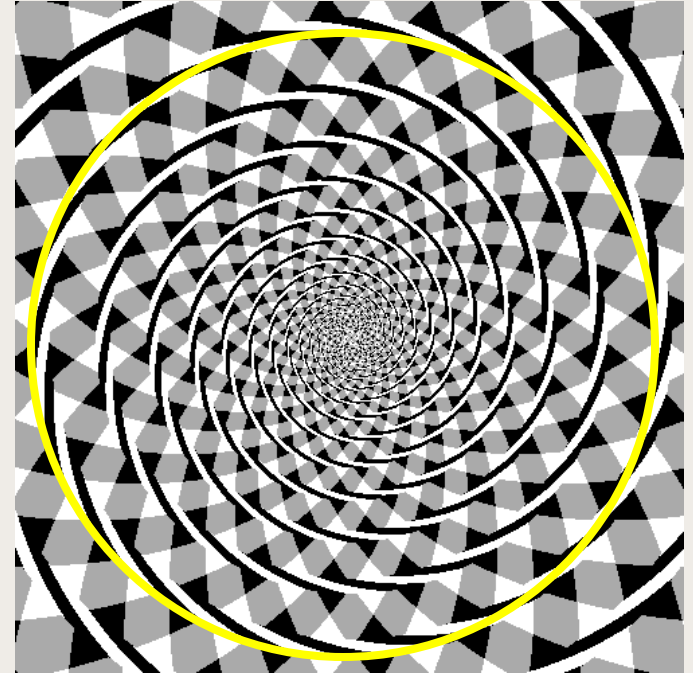
Trends in Theories of Knowledge Construction

- Decline of the dominance of the post-positivist paradigm
- Emergence of alternative epistemologies, especially constructivism
- Acceptance of qualitative methods
- Increased use of mixed-methods designs
- Increased interest in knowing more about programs than whether intervention x caused effect y .



Don Campbell's Critical Realism

- Posits the existence of a real world beyond the knower
- But perception of this world is distorted by visual and auditory illusions, cognitive schema and biases



Is this a spiral?

Don Campbell's Quest for Valid Inferences

- Threats to Valid Inferences
 - Threats to *internal validity* (causation)
 - History: events between pre- and posttest
 - Maturation: changes in participants over time
 - Testing: effects of taking test
 - Threats to *external validity* (generalizability)
 - Reactive effects

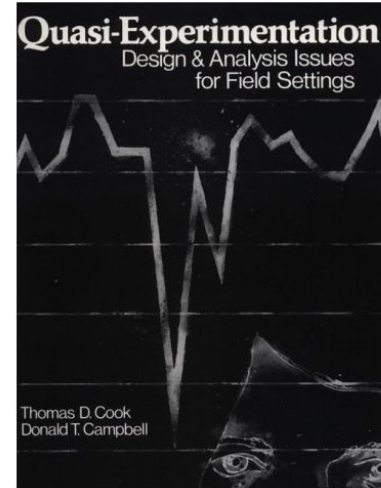


Campbell & Stanley, 1963



Quasi-Experimentation

- Quasi-experimental designs
 - Non-equivalent comparison groups
 - Time series
 - Regression discontinuity...



Cook & Campbell, 1979



Lee Cronbach stands Don Campbell on his head

- “all that can be validly meant by reference to a cause in a particular instance is that, on one trial of a partially specified manipulation of t under conditions A, B and C, along with other conditions not named, phenomena P was observed. To introduce the word *cause* seems pointless. Campbell’s writings make internal validity a property of trivial, past-tense, and local statements.” (Cronbach, 1982, p. 137)



Lee Cronbach and External Validity

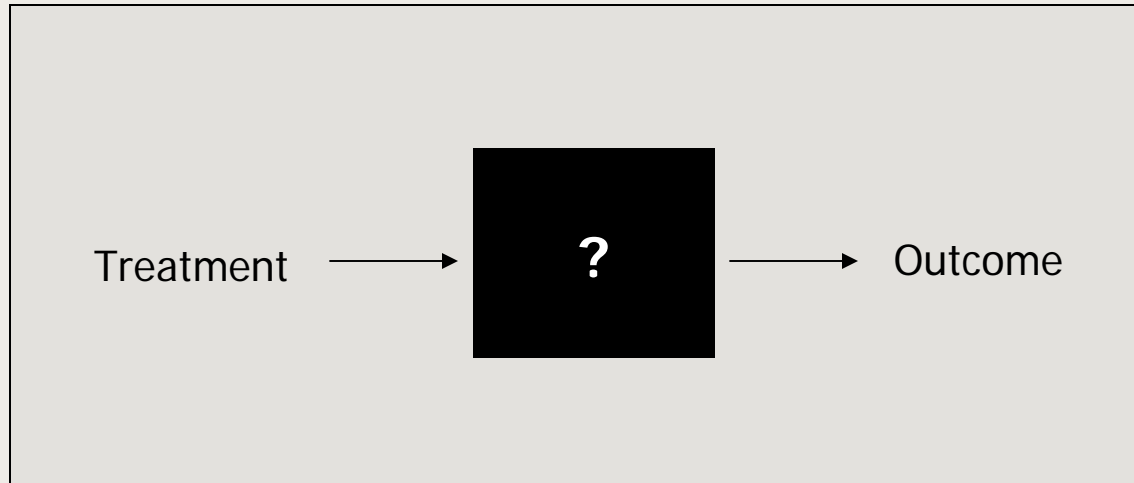
- What counts is *external validity*—the plausibility of conclusions about the *un-sampled* universe of persons, treatments, outcomes, settings.

UTOS → *UTOS

- UTOS (*sampled* units, treatments, outcomes, settings)
 - *UTOS (*un-sampled* units, treatments, outcomes, settings)
- Explaining *how* a treatment produces its effect promotes extrapolation to other populations.



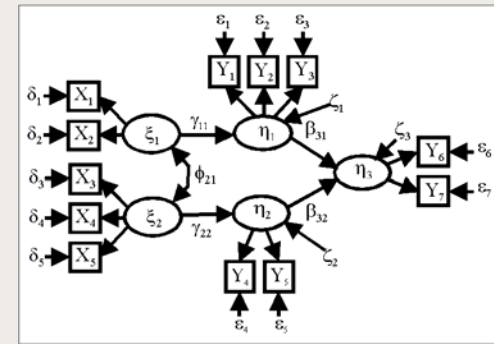
The Black Box of Experimentation



Inside the Black Box:

The theory-driven evaluations of Peter Rossi and Huey Chen

- TDE studies program inputs, causal mediating process, and program effects
- TDE uses structural equation modeling to reveal causal mediating processes between input and outputs
- Website: <http://www.mechanisms.org/>



Knowledge Construction

- Increased interest in understanding *how* interventions produce their effects so that these effects can be reproduced in other settings.



10 Minute Break



Activity 1

- Find someone from a different discipline
- Ask each other following question:
 - What are the typical assumptions in your discipline about ways of constructing credible knowledge about programs and their effects?
- Share what you learned about cross-disciplinary differences with the larger group



Enduring Issues in Evaluation Theory

Purpose of Evaluation

Social Programming

Knowledge Construction

➤ *Valuing*

Use

Practice



Theories of Valuing

- Key Issues
 - How can justifiable value statements relating to programs be constructed?
 - Whose values are considered in constructing such statements?



Theories of Valuing

- *Prescriptive theory*: theory that advocates the primacy of particular values
- *Descriptive theory*: theories that describe values without advocating one as best
- *Meta-theory*: the study of the nature of and justification for valuing



Scriven and the Science of Valuing

- Valuing as a science
 - “Evaluation research must produce as a conclusion exactly the kind of statement that social scientists for years have been taught is illegitimate: a judgment of value, worth, or merit” (Scriven, 1974, p.4)
- Values as scientific constructs
 - “Value, worth, quality, and merit are simply constructs from observable variables” (1986, p. 39).



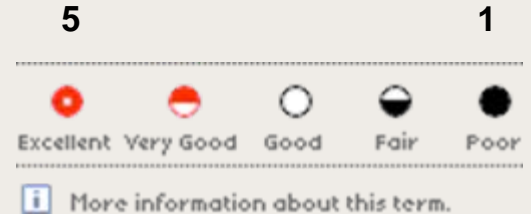
Scriven's Logic of Evaluation

1. Choose the object to be evaluated (the *evaluand*)
2. Establish criteria of merit
3. Establish standards of merit
4. Determine performance of the evaluand in relation to competitors or standards of merit
5. Summarize results into an evaluative judgment



The Logic of Valuing Vacuum Cleaners

CR Best Buy indicates an exceptional value



	Price	Overall Score					Test Results				
		P	F	G	VG	E	Carpet	Bare Floors	Noise	Emissions	Handling
Bissell 5770	249.90	[Red bar spanning P, F, G, VG, E]					5	4	5	5	5
Dirt Devil M110006	\$153.99	[Red bar spanning P, F, G, VG]					5	3	3	4	5
Dyson DC17	\$449.99	[Red bar spanning P, F, G, VG]					3	3	4	4	5
Hoover UH30010	\$299.99	[Red bar spanning P, F, G]					3	2	5	3	3

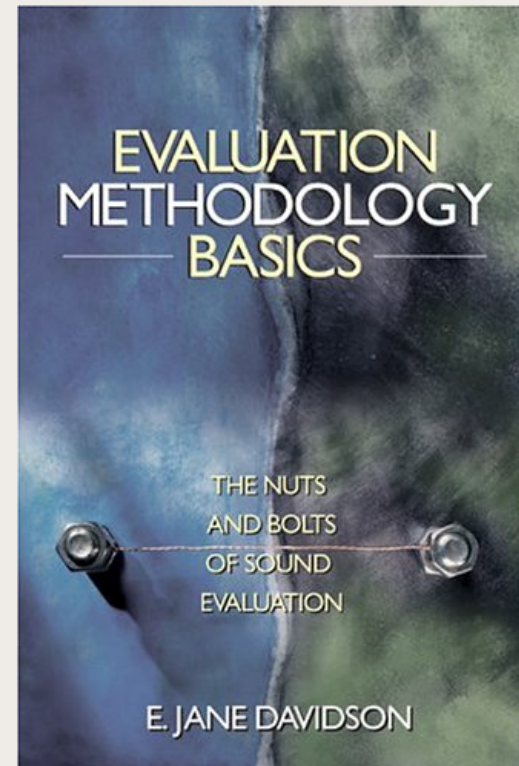
Evaluative Judgments

Criteria of Merit



The Logic of Evaluation

- A clear, concise explication of the fundamental logic of evaluation by one of Scriven's students.



Enduring Issues in Evaluation Theory

Purpose of Evaluation

Social Programming

Knowledge Construction

Valuing

➤ *Use*

Practice



Evaluation Use

- *How can evaluators produce results that are useful for social problem solving?*

Naïve Instrumentalism

- Evaluations often directly influence program decisions
- Evaluators don't need to make any special effort to ensure their findings are used



A multiplicity of uses and users

- Expanded typology of uses
- Identification of intended users and how they will make use of findings



Carol Weiss and Conceptual Use

- Instrumental use of evaluation findings is rare
- Conceptual use (“enlightenment”) is more common
 - “...diffuse and undirected infiltration of research ideas into their understanding of the world” (Weiss & Bucuvalas, 1980, p. 267).



Types of Evaluation Use

FINDINGS USE	PROCESS USE
Use of evaluation findings	Change that occurs as a result of participation in the evaluation process
•Instrumental	•Instrumental
•Conceptual	•Conceptual
•Legitimative	•Symbolic



-Alkin & Taut, 2003

The Personal Factor in Evaluation Use

- *The Personal Factor*: “the presence of an identifiable individual or group of people who personally care about the evaluation and the findings it generates” (Patton, 1997, p.44)
- Where the personal factor was present, the evaluation was used
- Findings led to Utilization-Focused Evaluation



Utilization-Focused Evaluation

Intended use by intended users

- Identify primary intended users
- Identify intended uses of findings
- Determine how findings might be used
- Given expected use, decide whether evaluation is worth doing
- If evaluation is worth doing, proceed with design and measurement decisions



Enduring Issues in Evaluation Theory

Purpose of Evaluation

Social Programming

Knowledge Construction

Valuing

Use

➤ *Practice*

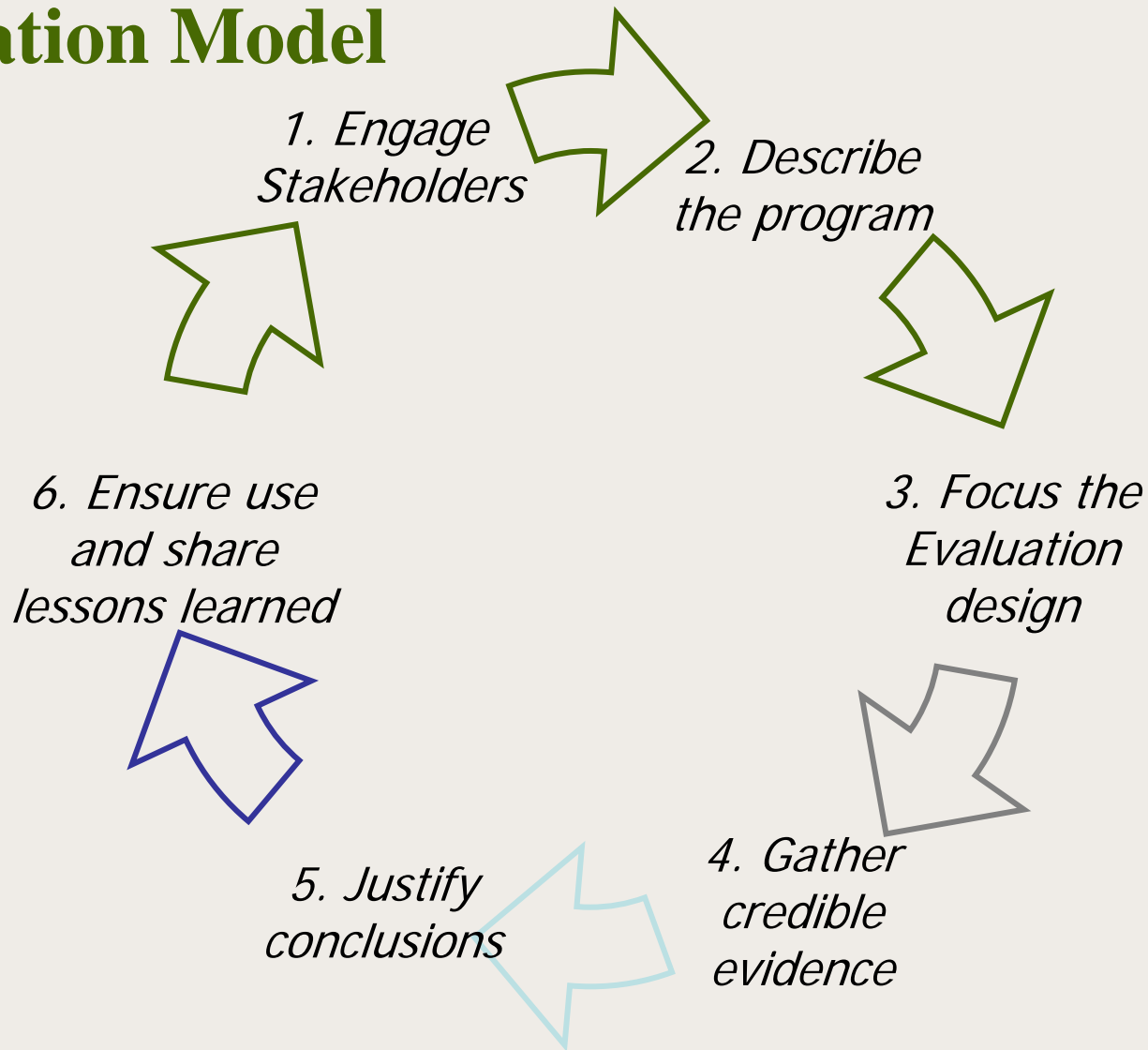


Evaluation Practice Issues

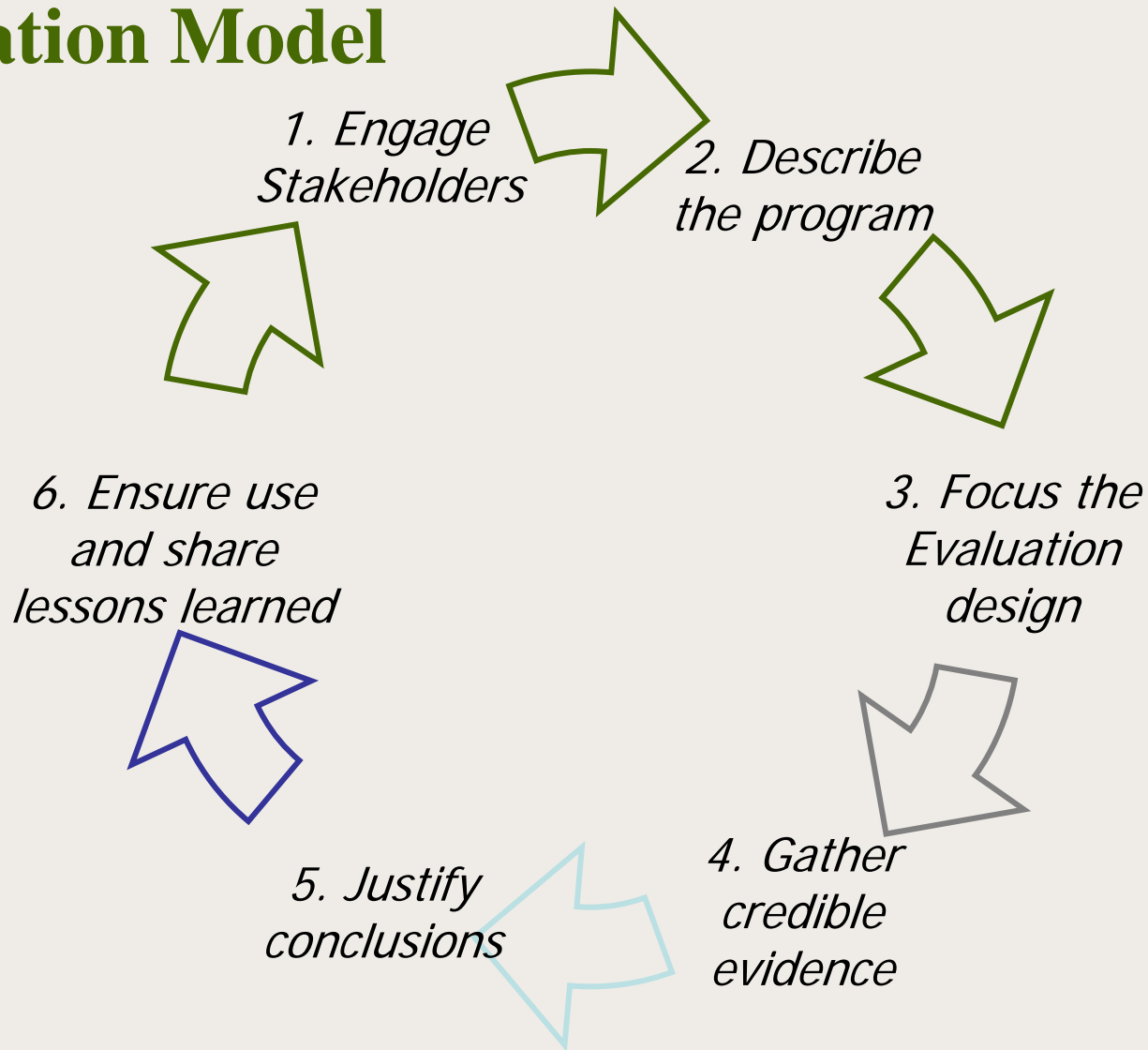
- When an evaluation should/should not be done
- Evaluator roles
- Stakeholder roles
- Types of questions
- Evaluation designs
- Ways to facilitate use



Evaluation Model



Evaluation Model



Summing Up

- Limits of typologies
- Need for a new field guide
- Enduring issues over time
- Toward more reflective evaluation practice



Activity 2

- Find a second person from a different discipline
- Ask each other following questions:
 - Do prescriptive or descriptive theories of valuing predominate in your discipline? Why?
- Share what you learned about cross-disciplinary differences with the larger group



References

- Alkin, M. (2004). *Evaluation roots: Tracing theorists' views and influences*. Thousand Oaks: Sage.
- Alkin, M. & Taut, S. (2003). Unbundling evaluation use. *Studies in Educational Evaluation*, 29, 1-12.
- Campbell, D. (1969). Reforms as Experiments. *American Psychologist*, 24, 409-429.
- Campbell, D. & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin Co.
- Cook, T. & Campbell, D. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin Co.
- Cronbach, L. (1982). *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.
- Davidson, E.J. (2004). *Evaluation methodology basics: The nuts and bolts of sound evaluation*. Thousand Oaks: Sage
- Greene, J.C. (2007). *Mixed methods in social inquiry*. San Francisco, CA: Jossey-Bass.
- Mark, M.M., Henry, G.T., & Julnes, G. (1999). Toward and integrative framework for evaluation practice. *American Journal of Evaluation*, 20(2), 177-198.
- Millstein, B., Wetterhall, S. (2000). A framework featuring steps and standards for program evaluation. *Health Promotion Practice*, 1(3), 221-228.
- Patton, M. (1997). *Utilization-focused evaluation: The new century text*. (3rd ed.) Thousand Oaks: Sage.



References

- Rossi, P. (1985). The iron law of evaluation and other metallic rules. Paper presented at State University of New York, Albany, Rockefeller College.
- Rossi, P., Freeman, H., & Lipsey, M. (1999). *Evaluation: A systematic approach* (6th ed.). Thousand Oaks: Sage.
- Scriven, Michael. (1967). The methodology of evaluation. In R.W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39-83). Chicago: Rand McNally.
- Scriven, M. (1974). Evaluation perspectives and procedures. In J.W. Popham (Ed.), *Evaluation in education: Current application* (pp. 3-9). Berkeley, CA: McCutchan.
- Scriven, M. (1986). New frontiers of evaluation. *Evaluation Practice*, 7, 7-44.
- Shadish, W., & Epstein, R. (1987). Patterns of program evaluation practice among members of Evaluation Research Society and Evaluation Network. *Evaluation Review*, 11, 555-590.
- Shadish, W., Cook, T., & Leviton, L. (1991). *Foundations of Program Evaluation: Theories of Practice*. Newbury Park: Sage.
- Stufflebeam, D., & Shinkfield, A. (2007). *Evaluation theory, models, and applications*. San Francisco: Jossey-Bass.
- Weiss, C. (1973). Where politics and evaluation research meet. *Evaluation*, 1, 37-45.
- Weiss, C. (1987). Evaluating social programs: What have we learned? *Society*, 25(1). 40-45.
- Weiss, C., & Bucuvalas, M. (1980). *Social science research and decision making*. New York: Columbia University Press.
- Weiss, C.H., Murphy-Graham, E., & Birkeland, S. (2005). An alternative route to policy influence: How evaluations affect D.A.R.E. *American Journal of Evaluation*, 26(1), 12-30.



Contact Information

University Outreach and Engagement

Michigan State University
Kellogg Center, Garden Level
East Lansing, MI 48824-1022
Phone: (517) 353-8977
Fax: (517) 432-9541
E-mail: outreach@msu.edu
Web site: outreach.msu.edu

