

# Model as a Concept in Evaluation Discourse

---

Frank G. Cookingham

*© World Vision International 2012, All Rights Reserved. No part of this document or any of its contents may be reproduced, copied, modified or adapted, without the prior written consent of the author.*

---

“Model” is a useful term when comparing different approaches to evaluation. But sometimes an author may be using the term differently than the meanings that the reader associates with the term. Consider the list of definitions for “model” in the appendix. Synonyms include: model, example, pattern, exemplar and ideal. The shared meaning is something set or held before someone for *guidance* or *imitation*.

This paper shares highlights of some evaluation models that are relevant for evaluation from a transformative perspective. A framework for an evaluation model is described that can show “at a glance” the distinguishing features of a particular approach to evaluation.

## Evaluation models classic text

*Evaluation Models* (Madaus, Scriven and Stufflebeam, 1983) was a classic text in the formative years of evaluation development. Ideas about models gleaned from the Preface (p. vi – viii).

- A model is a presentation that clarifies how central concepts are used and integrates them and their meanings into a coherent framework that guides all aspects of their work. A model is an evolving conceptualization of how theory and practice can guide practice in the field.
- A model is a school of thought by an acknowledged authority in the field.
- “[Each model in this volume] characterizes its author’s view of the main concepts involved in evaluation work and provides guidelines for using these concepts to arrive at defensible descriptions, judgments, and recommendations” (p. xiii). They are models for working in accord with a set of beliefs about evaluation.

About 20 years later this text was updated and expanded (Stufflebeam, Madaus and Kelleghan, 2000); I have not read it.

## Stufflebeam’s analysis of 22 models

Evaluation model → idealized view
--------------------------------------

Almost twenty years after the classic text was published Stufflebeam (2001) uses the term “evaluation approaches” instead of “models” in a comprehensive analysis. He does not use the term “model” because some of the approaches he examined are not legitimate for evaluation practice. In general “evaluation model” refers to an idealized

view of conducting evaluation. The purpose is to guide evaluation activities to achieve a specified purpose or set of objectives for the inquiry.

### **What is evaluation?**

Program evaluation for the purposes of Stufflebeam's monograph is characterized as "assessment of any set of coordinated activities directed at achieving goals" (2001, p. 10). In other words, the evaluand is a program defined as coordinated activities to achieve goals. The purpose of the assessment is captured in his definition of evaluation as "a study designed and conducted to assist some audience to assess an object's merit and worth" (2001, p. 11). This understanding of program evaluation is common in the literature.

Stufflebeam's list of 22 models was expanded to 26 approaches by Stufflebeam and Shinkfield (2007), but I have not read that book. Perhaps I will read it and revise this paper as needed. I am confident that this version of the paper is worth the time and effort it takes to understand it.

### **Four categories of approaches.**

Stufflebeam (2001) places the 22 approaches into four categories.

- Pseudoevaluations (n = 2). Two approaches promote invalid or incomplete findings often to achieve political objectives. Hence the label "psuedoevaluations."
- Questions or Methods-Oriented Evaluations (n = 13). These approaches come out of the view of evaluation as a form of scientific research.
- Improvement/Accountability Evaluations (n = 3). This group of evaluations provides comprehensive trustworthy information for assessing merit and worth.
- Social Agenda/Advocacy Evaluations (n = 4). In this group evaluators see evaluation as a process for making a difference in society. Evaluation is done to empower disenfranchised groups.

Ten descriptors are used to differentiate the 22 approaches. This paper briefly describes key features of the two pseudoevaluations and four social agenda evaluations. I have drawn from Stufflebeam's descriptions extensively but selectively. This is the information most relevant for planning program evaluation from a transformative perspective.

### **Pseudoevaluations.**

These two approaches are not evaluations because they violate the fundamental characteristic that the conclusions of the inquiry must be trustworthy for any unbiased person that wants to use them for making sound decisions.

- Public-Relations Inspired Studies (advertising, infomercial). The goal of the exercise is to provide persuasive information to constituents that a program is

- sound and effective. Typically no information is provided about weaknesses or limitations, and information about strengths may be exaggerated.
- **Politically Controlled Studies**. Typically the goal of the exercise is accurate information about the object of study. The client, however, may not release findings to those who have a legitimate right to them, or may release only some findings, or may misrepresent the findings to support the client's ambitions.

### **Social Agenda/Advocacy Evaluations.**

Except for Patton's utilization model, evaluation exercises are designed to improve some aspect of society. In general they seek to ensure that all segments of society have equal access to opportunities and services. They give preferential treatment to the disadvantaged. They seek to empower the disenfranchised. Patton's model is included in this group because it emphasizes intense stakeholder participation throughout the exercise, and considerable control over what is evaluated and how it is given to the user group.

Philosophical underpinnings include cultural pluralism, moral relativity and multiple realities. They provide for democratic engagement of stakeholders in obtaining and interpreting findings. Adequate safeguards need to be in place so that the inquiry meets evaluation standards.

Stufflebeam included all four approaches in his list of nine approaches most worthy of being used. Note that Mertens' (2009) work was not reviewed by Stufflebeam. She began publishing articles soon after Stufflebeam wrote the monograph. There are substantive similarities between these approaches and Mertens' approach.

### ***Client-Centered Studies (Responsive Evaluation).***

Robert Stake (1983) may be the evaluator most associated with this approach.

In essence this approach formalizes informal intuitive evaluation. The exercise is an ongoing conversation between the evaluator and clients focused on discovering, investigating and addressing program issues. A key product of the evaluation is a rich description of various aspects of the program. Special attention is given to side effects, different standards that different people have for the program and their judgments of the program.

### ***Action research strength.***

The action research approach combined with consideration of how experts judge the program is a primary strength of this approach. Participation in a client-centered study is likely to evaluation a variety of evaluation skills that can be applied elsewhere.

The evaluator interacts continuously with program clients and other stakeholders to identify and respond to their evaluation needs. The evaluator assists the people involved in a program to evaluate it, and then improve it. The evaluator does not seek an overall judgment of merit and worth. Underlying philosophical positions that guide the evaluator:

- Promote equity and fairness
- Help those with little power; thwart the misuse of power
- Expose the huckster (dig deeply into the program rationale to see if it is based on tested knowledge)
- Unnerve those who believe they understand how the program operates and what it achieves
- Reassure those who are insecure about program operations and results
- Always help people see things from alternative viewpoints

### *Weaknesses.*

A major weakness is that the final product may not be regarded as credible in other settings since the program stakeholders influenced the content of the inquiry. This approach is not helpful for accountability purposes. Sometimes the evaluation may exacerbate existing conflicts among different groups rather than identify common ground for moving forward.

### ***Constructivist Evaluation.***

Yvonne Lincoln and Egon Guba are the pioneers of this approach (Lincoln and Guba 1985, Guba and Lincoln 1989).

This approach to evaluation evolved out of strong rejection of the scientific worldview based on the fundamental assumption that there is a reality that is not influenced by people studying the reality, and that people using appropriate techniques of inquiry can understand that reality. The main purpose of the exercise is to determine and make sense of the relevant cognitive constructions that different stakeholders have for whatever is being evaluated. The evaluator is expected to empower those who are disenfranchised while respecting the free will of all participants to shape the exercise as it proceeds.

The evaluation questions emerge as the evaluator and stakeholders engage in conversations and hermeneutic exercises about their views about the program. The set of questions is never considered fixed; it changes as information is collected and processed. In the early stages of the exercise stakeholders are encouraged to questions and what those questions mean to them, regardless of what others have shared. Then the evaluator facilitates a dialectal process toward consensus, reminding participants

that consensus is not required but that exploration can be informative. This exploration of consensus is unique among the four social agenda evaluation models.

In this approach all cognitive constructions remain open to revision as the exercise proceeds. No construction is deemed truer than any other, but stakeholder views may converge on some of them as more informed than others. Knowledge is viewed as constantly evolving; there are no right answers to evaluation questions.

Fetterman's (2001) empowerment evaluation is similar to the constructivist approach in many ways. A key difference is that Fetterman takes a back seat in the process, acting only as a technical advisor. Control of the direction taken by the exercise is in the hands of the stakeholders. Constructivist evaluators retain control of the process, but work hard to be inclusive and ethical in their decisions about the direction of the exercise.

#### *Advantages.*

- Stakeholders who engage fully throughout the exercise will become more informed about the program, sometimes in unexpected ways. Also, they are more likely to be involved in post-evaluation change efforts.
- Qualitative research approaches and triangulation provide rich descriptions of those aspects of the program that are studied.
- Participants will become more aware of the role of perception in their own thinking and in the thinking of others.

#### *Disadvantages.*

- This approach will not provide adequate evidence for accountability. If this is what stakeholders want, this approach should not be implemented.
- Many stakeholders will reject the open-ended nature of the exercise and the extensive involvement required of all stakeholders throughout the exercise. To be effective, everyone must be fully informed about the approach and genuinely agreed to participate fully.
- Turnover in stakeholders is a major limitation, as it may take considerable time and effort to bring new participants up to speed. This will try the patience of other stakeholders, and the evaluator.
- Stakeholders who are poorly informed about the issues being studied will hinder the process. It may be unrealistic for the evaluator to take the time and effort to prepare them to participate meaningfully in the process.
- Some stakeholders may not tolerate an evaluator who refuses to declare which constructions are better than others, or to offer "hard" data to resolve differences among the constructions.
- If consensus is reached, which is not always the case, it cannot be applied to other settings.

### ***Deliberative Democratic Evaluation.***

Ernest House and Kenneth Howe (2000) are primary advocates for this approach. It has three main dimensions that are applied in every stage of the evaluation process.

- Democratic participation by a representative group of stakeholders. Equity is a major focus; power imbalances are not tolerated. Fair representation of all legitimate views is a major focus. The evaluator determines the final questions that guide the evaluation after sufficient dialogue and deliberation.
- Dialogue to examine and authenticate stakeholders' inputs.
- Deliberation to arrive at a defensible assessment of the program's merit and worth. The evaluator prepares the assessment after seriously reflecting on responses by stakeholders to preliminary findings. The final assessment is not determined by majority vote of stakeholders.

This model is unique in this group in that the evaluator makes the final judgments, not the stakeholders. But it is an ideal model that is very difficult to implement. It requires extensive participation by a large number of people that are willing to be open to different views of what is important, and different hierarchies of values. It requires participants to trust the evaluator in making final judgments fairly and to respect those judgments where they differ from their own conclusions.

### ***Utilization-Focused Evaluation.***

Michael Quinn Patton (1986) is the evaluator most associated with this approach. It has had a major influence on evaluation work and it has inspired considerable research on use of evaluation findings. See the fourth edition (2008) for in-depth discussion. (I have not read the 3<sup>rd</sup> edition published in 1997 nor the 4<sup>th</sup> edition.)

The purpose of this approach is to identify what information different stakeholder groups need, and how they will use it. Intense interaction between a group of users and the evaluator occurs throughout the process. As a consequence, the focus of the evaluation may shift as preliminary findings are considered.

The criterion for success in this approach is program improvement by application of evaluation findings. The evaluator is a servant of information users; they determine what information is needed, how and when it will be reported, and how it will be used. All legitimate evaluation methods can be used; the evaluator negotiates the ways that will meet the users' information needs and meet professional evaluation standards.

This approach does not necessarily advocate for a particular social agenda.

### *Strength.*

The primary strength of this approach is that evaluation findings actually make a difference in whatever is evaluated. Generally target users want to know about program impacts, costs, and cost-benefits.

### *Limitations.*

- Turnover in the group of targeted users can disrupt the exercise greatly. In a major exercise turnover is inevitable.
- The evaluator must be competent in all evaluation methodologies and a skilled negotiator. He must understand all the professional standards and be able to satisfy them creatively. The pool of evaluators that can lead major exercises is relatively small.
- Some members of the user group can corrupt the process if they have power over other members.

## **Using a framework to describe an evaluation model**

The concepts of framework, approach (Stufflebeam, 2001) or model (Madaus, Scriven and Stufflebeam, 1983), and design are hierarchical. Different evaluators may describe them differently; I find it helpful to see the hierarchy moving from abstract framework to generic model to a specific design (evaluation plan).

### **Framework.**

A framework is a conceptual structure for organizing ideas. The five fundamental issues that undergird practical program evaluation (Shadish, Cook and Leviton, 1991) is an example of a framework. The five issues are (p.35):

- Social programming – What are the important problems this program could address? Can the program be improved? Is it worth doing so? [If it is not worth improving], what is worth doing from an evaluation perspective?
- Knowledge use – How can I make sure my results get used quickly to help this program? Do I want to focus on rapid use of results? If not, can my evaluation be useful in other ways?
- Valuing – Is this a good program? By which notion of “good”? What justifies the conclusion about the goodness of the program?
- Knowledge construction – How do I know all this? What counts as a confident answer to such a question? What causes that confidence?
- Evaluation practice – Given my [limits], and given the seemingly unlimited possibilities, how can I narrow my options to do a feasible evaluation? What is my role [worth] as educator, methodological expert, or judge of program? What questions should I ask, and what methods should I use [to answer them, or to discover other important questions]?

Other frameworks can be structured around approaches to inquiry (case study, experimental, survey, focus groups, etc.) or purpose of evaluation (add to a body of knowledge, determine attribution, demonstrate accountability for achieving goals, etc.).

### **Model.**

A model is the detailed description of an evaluation approach to dealing with the five issues. See Stufflebeam (2001) for examples of a variety of models and then describe those of interest in terms of these five fundamental issues.

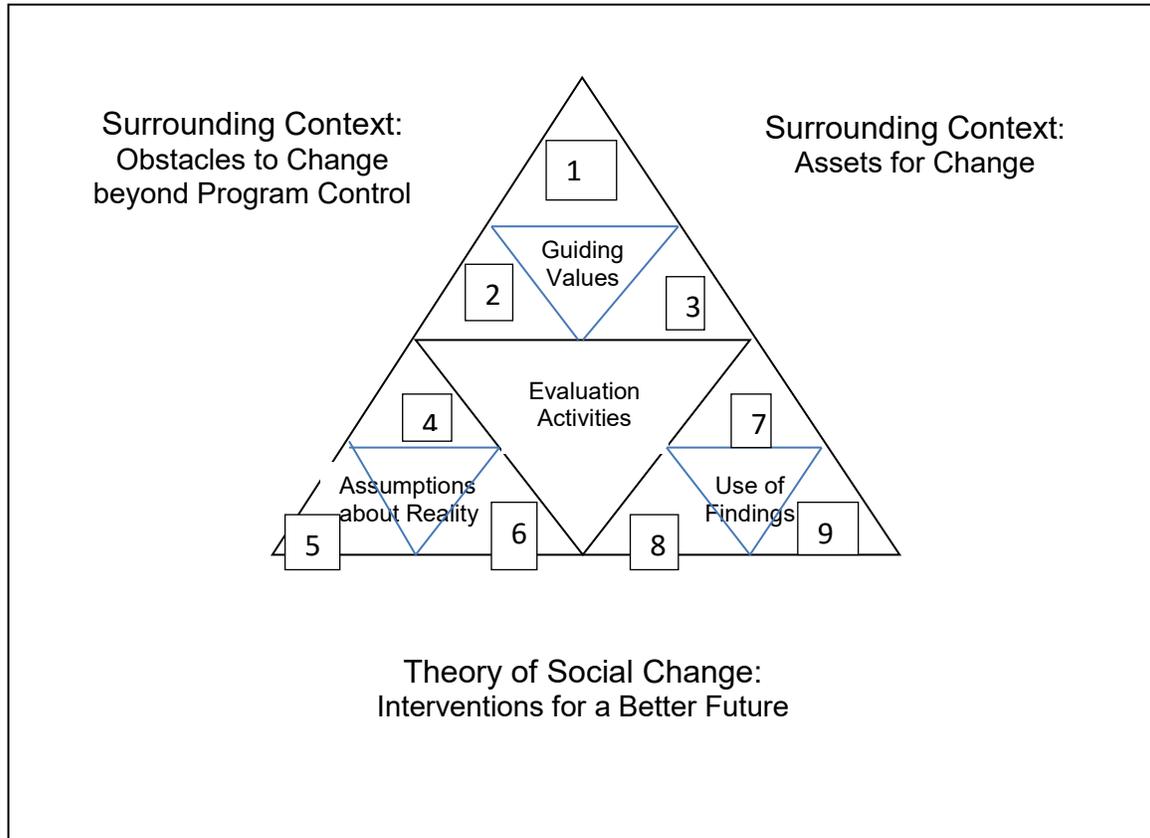
### **Design.**

This is a description of what data will be collected, how it will be collected, and how it will be analyzed as a particular model, or combination of models, is applied to the evaluand.

### **Application.**

I have organized the evaluation framework described above into a graphic that can show at a glance the defining features of an evaluation model.

## General Framework for an Evaluation Model



The General Framework simply reorganizes the five components described above to show that evaluation activities (the center of the framework) are determined by the evaluator's understanding of:

- Surrounding context within which the program is embedded and the underlying assumptions that support the implementation of the program (theory of social change). Two clusters of factors in the context are the obstacles to change (risks) and the assets for change in the program setting. This is the domain of the social programming component.
- Values that guide interpreting evidence to conclude to what extent the program is good and bad, helpful to some stakeholders and harmful to others, worth the cost, etc.
- Assumptions about reality and how you know what is real is the essence of the knowledge construction component. Different sets of assumptions in this area define very different evaluation activities.
- Utilization factors that influence how different groups will interpret and use conclusions and recommendations.

The purpose of having three aspects for guiding values, assumptions about reality and use of findings is to provide a visual overview of the defining features for a model.

There is nothing magical about three aspects per supporting component. The intellectual effort to identify the nine aspects that influence evaluation activities may help one understand what differentiates one model from others beyond the particular terminology that advocates use. In a particular case seven aspects may be sufficient while in another case twelve aspects may be necessary.

## References

- Fetterman, David M. (2001). *Foundations of empowerment evaluation*. Thousand Oaks, California USA: Sage Publications.
- Guba, Egon G., and Lincoln, Yvonne S. (1989). *Fourth generation evaluation*. Thousand Oaks, California USA: Sage Publications.
- House, Ernest R., and Howe, Kenneth R. (2000). Deliberative democratic evaluation. In Katherine E. Ryan, and Lizanne DeStefano (Eds.), *Evaluation as a democratic process: Promoting inclusion, dialogue, and deliberation* (pp. 3-12). *New directions for evaluation, 85*. San Francisco, California USA: Jossey-Bass.
- Lincoln, Yvonne S., and Guba, Egon G. (1985). *Naturalistic inquiry*. Thousand Oaks, California USA: Sage Publications.
- Madaus, George. F., Scriven, Michael, & Stufflebeam, Daniel L. (Eds.). (1983). *Evaluation models: Viewpoints on educational and human services evaluation*. Boston, Massachusetts USA: Kluwer-Nijoff.
- Mertens, Donna M. (2009). *Transformative research and evaluation*. New York, New York: The Guildford Press.
- Patton, Michael Quinn. (1986). *Utilization-focused evaluation* (2<sup>nd</sup> ed.). Thousand Oaks, California USA: Sage Publications.
- Shadish, William R., Cook, Thomas D., & Leviton, Laura C. (1991). *Foundations of program evaluation: Theories of practice*. Thousand Oaks, California USA: Sage Publications.
- Stake, Robert E. (1983). Program evaluation, particularly responsive evaluation. In George F. Madaus, Michael Scriven, & Daniel L. Stufflebeam (Eds.). (1983). *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 287-310). Boston, Massachusetts USA: Kluwer-Nijoff.
- Stufflebeam, Daniel L. (2001). *Evaluation models. New Directions for Evaluation, 89*. San Francisco, California Spring 2001

Stufflebeam, Daniel L., Madaus, George F., & Kelleghan, Thomas. (Eds.). (2000). Evaluation models: Viewpoints on educational and human services evaluation (2<sup>nd</sup> ed.). Boston, Massachusetts USA: Kluwer-Nijoff.

Stufflebeam, Daniel L., and Shinkfield, Anthony J. (2007). *Evaluation theory, models, & applications*. San Francisco, California USA: Jossey-Bass.

### **Appendix: Senses of the noun “model”**

These entries are from Webster’s New Collegiate Dictionary (8<sup>th</sup> edition), 1977.

- A miniature representation of something, as in “a model train was set up on the floor of the children’s bedroom.”
- A copy or image of something, as in “John’s photograph was used by the artist as a model for painting his portrait.”
- A person or thing that serves as a pattern for an artist, as in “Mary posed as a model for the students in art class because she knew how much life drawing had helped her develop as an artist.”
- A pattern for something, as in “Susan made a beautiful dress by following the pattern that she bought at the fabric store.”
- An example for imitation or emulation, as in “a fashion model inspires many women to change their hair style.”
- An archetype, or ideal representation of something.
- A person who is employed to display clothes or merchandise like a mannequin.
- A type or design of something, as in “the newest car models are described in this issue of the magazine.”
- A description or analogy that is used to visualize something that cannot be observed directly, such as an atom.
- A system of postulates, data and inferences presented as a mathematical description of an entity or state of affairs, as in “mathematical model for the trajectory of a rocket.”