
Evaluation Models

Evaluation in Education and Human Services

Editors:

George F. Madaus, Boston College, Chestnut Hill, Massachusetts, U.S.A.
Daniel L. Stufflebeam, Western Michigan University, Kalamazoo, Michigan, U.S.A.

Previously published books in the series:

Kellaghan, T., Madaus, G., and Airasian, P.:
The Effects of Standardized Testing
Madaus, G. (editor):
The Courts, Validity, and Minimum Competency Testing
Brinkerhoff, R., Brethower, D., Hluchyj, T., and Nowakowski, J.:
Program Evaluation, Sourcebook/Casebook
Brinkerhoff, R., Brethower, D., Hluchyj, T., and Nowakowski, J.:
Program Evaluation, Sourcebook
Brinkerhoff, R., Brethower, D., Hluchyj, T., and Nowakowski, J.:
Program Evaluation, Design Manual

Evaluation Models

*Viewpoints on Educational and
Human Services Evaluation*

**George F. Madaus
Michael S. Scriven
Daniel L. Stufflebeam**



Kluwer-Nijhoff Publishing
Boston The Hague Dordrecht Lancaster

Distributors for North America:

Kluwer Academic Publishers
101 Philip Drive
Assinippi Park
Norwell, MA 02061

Distributors outside North America:

Kluwer Academic Publishers Group
Distribution Centre
P.O. Box 322
3300AH Dordrecht,
THE NETHERLANDS

Library of Congress Cataloging in Publication Data

Main entry under title:

Evaluation models.

(Evaluation in education and human services)
includes index.

1. Educational surveys — Addresses, essays, lectures.
2. School management and organization — Evaluation — Addresses, essays, lectures.
3. Curriculum evaluation — Addresses, essays, lectures. I. Madaus, George F. II. Scriven, Michael. III. Stufflebeam, Daniel L. IV. Series. LB2823.E87 1983 379.1'54 82-22562

ISBN-13: 978-94-009-6671-0 e-ISBN-13: 978-94-009-6669-7
DOI: 10.1007/978-94-009-6669-7

Copyright © 1983 by Kluwer-Nijhoff Publishing

Softcover reprint of the hardcover 1st edition 1983

Tenth Printing, 1993.

No part of this book may be reproduced in any form by print, photoprint, microfilm, or any other means without written permission of the publisher.

Contents

List of Figures and Tables	ix
Preface	xi
I	
An Overview of Models and Conceptualizations	1
1	
Program Evaluation: A Historical Overview	
George F. Madaus, Daniel L. Stufflebeam, and Michael S. Scriven	3
2	
An Analysis of Alternative Approaches to Evaluation	
Daniel L. Stufflebeam and William J. Webster	23
3	
Assumptions Underlying Evaluation Models	
Ernest R. House	45
II	
Models and Conceptualizations	65
4	
A Rationale for Program Evaluation	
Ralph W. Tyler	67
5	
The Discrepancy Evaluation Model	
Andrés Steinmetz	79
	v

6		
	Course Improvement through Evaluation	
	Lee J. Cronbach	101
7		
	The CIPP Model for Program Evaluation	
	Daniel L. Stufflebeam	117
8		
	The Evaluation of Broad-Aim Programs: Experimental Design, Its Difficulties, and an Alternative	
	Robert S. Weiss and Martin Rein	143
9		
	Societal Experimentation	
	Peter W. Airasian	163
10		
	Rationality to Ritual: The Multiple Roles of Evaluation in Governmental Processes	
	Robert E. Floden and Stephen S. Weiner	177
11		
	The Use of Judicial Evaluation Methods in the Formulation of Educational Policy	
	Robert L. Wolf	189
12		
	Deep Dark Deficits of the Adversary Evaluation Model	
	W. James Popham and Dale Carlson	205
13		
	The Clarification Hearing: A Personal View of the Process	
	George F. Madaus	215
14		
	Evaluation Ideologies	
	Michael S. Scriven	229
15		
	Flexner, Accreditation, and Evaluation	
	Robert E. Floden	261
16		
	The Case Study Method in Social Inquiry	
	Robert E. Stake	279

17		
Program Evaluation, Particularly Responsive Evaluation		
Robert E. Stake		287
18		
Epistemological and Methodological Bases of Naturalistic Inquiry		
Egon G. Guba and Yvonna S. Lincoln		311
19		
Educational Connoisseurship and Criticism: Their Form and Functions in Educational Evaluation		
Elliot W. Eisner		335
20		
The Explication Model: An Anthropological Approach to Program Evaluation		
Kent L. Koppelman		349
21		
Designing Evaluations of Educational and Social Progress by Lee J. Cronbach: A Synopsis		
Anthony J. Shrinkfield		357
22		
The Progress of Educational Evaluation: Rounding the First Bends in the River		
Nick L. Smith		381
III		
The Standards and the Ninety-five Theses		393
23		
The Standards for Evaluation of Educational Programs, Projects, and Materials: A Description and Summary		
Daniel L. Stufflebeam and George F. Madaus		395
24		
Ninety-five Theses for Reforming Program Evaluation		
Lee J. Cronbach		405
Index		413
Contributing Authors		423

List of Figures and Tables

List of Figures

Figure 3–1. A Scheme Relating Major Evaluation Models to the Philosophy of Liberalism	49
Figure 5–1. Level I Analysis	90
Figure 5–2. Level II Component Analysis	90
Figure 5–3. Level III Component Analysis of the Teacher In-Service Program	92
Figure 5–4. Level III Program Network	92
Figure 7–1. A Flowchart Depicting the Role of CIPP Evaluation in Effecting System Improvement	125
Figure 17–1. A Layout of Statements and Data to Be Collected by the Evaluator of an Educational Program	295
Figure 17–2. Prominent Events in a Responsive Evaluation	296
Figure 20–1. The Explication Model: A Teacher-Centered Evaluation	352

List of Tables

Table 2–1. An Analysis of Political-Orientation Study Types (Pseudo-Evaluation)	37
Table 2–2. An Analysis of Questions-Orientation Study Types (Quasi-Evaluation)	38
Table 2–3. An Analysis of Values-Orientation Study Types (True Evaluation)	39
Table 3–1. A Taxonomy of Major Evaluation Models	48
Table 5–1. First Attempt at a Standard	83

Table 5–2. Elements of an Evaluation Workplan	84
Table 5–3. Input-Process-Output Description for Two Components	91
Table 5–4. Kinds of Evaluation Useful to Program Improvement	94
Table 5–5. Summary of Data Collection Plan	96
Table 5–6. Summary of Data Collection Plan (continued)	97
Table 7–1. The Relevance of Four Evaluation Types to Decision Making and Accountability	125
Table 7–2. Four Types of Evaluation	129
Table 7–3. Outline for Documenting Evaluation Designs	138
Table 17–1. Comparison of Preordinate and Responsive Evaluators	299
Table 17–2. Nine Approaches to Evaluation	304
Table 18–1. Axiomatic Differences Between the Rationalistic and Naturalistic Paradigms	315
Table 23–1. Summary of the Standards for Evaluation of Educational Programs, Projects, and Materials	399
Table 23–2. Analysis of the Relative Importance of 30 Standards in Performing 10 Tasks in an Evaluation	404

Preface

Attempting formally to evaluate something involves the evaluator coming to grips with a number of abstract concepts such as value, merit, worth, growth, criteria, standards, objectives, needs, norms, client, audience, validity, reliability, objectivity, practical significance, accountability, improvement, process, product, formative, summative, costs, impact, information, credibility, and — of course — with the term *evaluation* itself. To communicate with colleagues and clients, evaluators need to clarify what they mean when they use such terms to denote important concepts central to their work. Moreover, evaluators need to integrate these concepts and their meanings into a coherent framework that guides all aspects of their work. If evaluation is to lay claim to the mantle of a profession, then these conceptualizations of evaluation must lead to the conduct of defensible evaluations.

The conceptualization of evaluation can never be a one-time activity nor can any conceptualization be static. Conceptualizations that guide evaluation work must keep pace with the growth of theory and practice in the field. Further, the design and conduct of any particular study involves a good deal of localized conceptualization. In any specific situation, the evaluator needs to define and clarify for others the following: the audiences and information requirements, the particular object to be evaluated, the purposes of the study, the inquiry approach to be employed, the concerns and issues to be examined, the variables to be assessed, the bases for interpreting findings, the communication mode to be used, the anticipated uses of the findings, and the standards to be invoked in assessing the quality of the work.

It is a small wonder, then, that attempts to conceptualize evaluation have been among the most influential works in the fast-growing literature of evaluation. The contents of this anthology attest to the fact that there has been a rich array of theoretical perspectives on evaluation. Given the complexity of evaluation work, the wide range of evaluative situations, the political contexts within which studies

occur, the service orientation of evaluations, and the varied backgrounds and beliefs of those who write about evaluation, it is easy to understand why the various generalized conceptualizations of evaluation found in the literature differ in many important respects. The ways that evaluation is conceptualized differ over the role of objectives in the process, the desirability of presenting convergent or divergent findings, the use or absence of experimental controls, and the place of hard or soft data in arriving at conclusions. It is also understandable that given evaluators sometimes follow one general approach in one kind of evaluation assignment and a quite different approach in another setting. Since the contexts in which evaluations take place are so variable, it is fortunate that evaluators can look to the literature for optional ways to conceptualize the evaluation process in order to find one which best suits a particular context.

Amidst this diversity of conceptual approaches to evaluation, however, a consensus has begun to emerge regarding the principles that should undergird all evaluations. This consensus is embodied in the two major sets of standards for evaluations that have been issued recently by the Joint Committee on Standards for Educational Evaluation and by the Evaluation Research Society. The appearance of these standards is one sign — but a sure one — that evaluation has begun to mature as a profession.

This book is an up-to-date reflection of the conceptual development of evaluation, particularly program evaluation, and is divided into three major sections. The first includes a historical perspective on the growth of evaluation theory and practice and two comparative analyses of the various alternative perspectives on evaluation. The second part contains articles that represent the current major schools of thought about evaluation, written by leading authors in the field of evaluation, including, articles by Tyler, Scriven, Stake, Eisner, Floden, Airasian, Guba and Lincoln, Stufflebeam, Cronbach, Steinmetz (on Provus's work), Weiss and Rein, Madaus, and Koppelman. These articles cover objectives-oriented evaluation, responsive evaluation, consumer-oriented evaluation, decision and improvement-oriented evaluation, naturalistic evaluation, discrepancy evaluation, adversarial evaluation, connoisseur evaluation, accreditation, accountability, and social experimentation. This section concludes with a forecast on the future of evaluation by Nick Smith. The final section describes and discusses the recently released *Standards for Evaluations of Educational Programs, Projects, and Materials* and summarizes the 95 theses recently issued by Cronbach and Associates in calling for a reformation of program evaluation.

In one sense, the core of this book presents a set of alternative evaluation models. These are not models in the sense of mathematical models used to test given theories, but they are models in the sense that each one 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. We are aware that some writers in the field have urged against according alternative perspectives on evaluation the status of models; but we think the alternative suggestion that these alternatives be called something else, such as persuasions or beliefs, would do little more than puzzle the readers. We are comfortable in presenting the alternative conceptualizations of evaluation that appear in the second part of the book, not as models *of* evaluation as it does occur, but as models *for* conducting studies according to the beliefs about evaluation that are held by the various authors. In this sense, they are idealized or “model” views of how to sort out and address the problems encountered in conducting evaluations.

We owe an enormous debt to the authors of the articles that appear in this book. We would like also to thank the various journals that gave us permission to reprint key pieces. We especially wish to thank Ralph Tyler and Peter Airasian for writing articles specifically for this book, as well as Egon Guba and Yvonna Lincoln, who adapted their article to fit within our space limitations. We also are grateful to Phil Jones, our publisher, who consistently supported our developmental effort. Thanks is extended also to Carol Marie Spoelman, Caroline Pike, and Mary Campbell for their competent clerical assistance. Special thanks to Rita Comtois for her administrative assistance throughout the project. Bernie Richey’s editorial help throughout is appreciated.

We believe this book should be of interest and assistance to the full range of persons who are part of any evaluation effort, including, especially, the clients who commission evaluation studies and use their results, evaluators, and administrators and staff in the programs that are evaluated. We believe the book should be useful as a text for courses in program evaluation and for workshops as well. Further, it should prove to be an invaluable reference book for those who participate in any aspect of formal evaluation work. We hope that this book will assist significantly those involved in program evaluation to increase their awareness of the complexity of evaluation; to increase their appreciation of alternative points of view; to improve their ability to use theoretical suggestions that appear in the literature; to increase their testing and critical appraisal of the various approaches; and, ultimately, to improve the quality and utility of their evaluations.