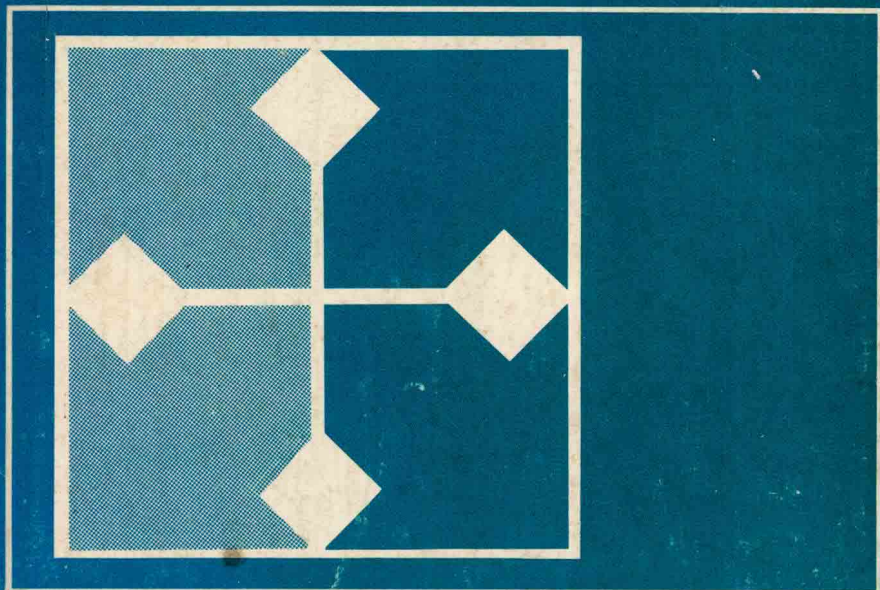


IMPROVING EVALUATIONS

EDITED BY
LOIS-ELLIN DATTA AND
ROBERT PERLOFF



A SAGE FOCUS EDITION

IMPROVING EVALUATIONS

SAGE FOCUS EDITIONS

1. **POLICE AND SOCIETY**
edited by DAVID H. BAYLEY
2. **WHY NATIONS ACT: THEORETICAL PERSPECTIVES FOR COMPARATIVE FOREIGN POLICY STUDIES**
edited by MAURICE A. EAST, STEPHEN A. SALMORE, and CHARLES F. HERMANN
3. **EVALUATION RESEARCH METHODS: A BASIC GUIDE**
edited by LEONARD RUTMAN
4. **SOCIAL SCIENTISTS AS ADVOCATES: VIEWS FROM THE APPLIED DISCIPLINES**
edited by GEORGE H. WEBER and GEORGE J. McCALL
5. **DYNAMICS OF GROUP DECISIONS**
edited by HERMANN BRANDSTATTER, JAMES H. DAVIS, and HEINZ SCHULER
6. **NATURAL ORDER: HISTORICAL STUDIES OF SCIENTIFIC CULTURE**
edited by BARRY BARNES and STEVEN SHAPIN
7. **CONFLICT AND CONTROL: CHALLENGE TO LEGITIMACY OF MODERN GOVERNMENTS**
edited by ARTHUR J. VIDICH and RONALD M. GLASSMAN
8. **CONTROVERSY: POLITICS OF TECHNICAL DECISIONS**
edited by DOROTHY NELKIN
9. **BATTERED WOMEN**
edited by DONNA M. MOORE
10. **CRIMINOLOGY: NEW CONCERNS**
edited by EDWARD SAGARIN
11. **TO AUGUR WELL: EARLY WARNING INDICATORS IN WORLD POLITICS**
edited by J. DAVID SINGER and MICHAEL D. WALLACE
12. **IMPROVING EVALUATIONS**
edited by LOIS-ELLIN DATTA and ROBERT PERLOFF

IMPROVING EVALUATIONS

EDITED BY
LOIS-ELLIN DATTA AND
ROBERT PERLOFF



SAGE Publications

Beverly Hills/London

Copyright © 1979 by Sage Publications, Inc.

All rights reserved. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

Royalties from the sales of this book are being donated to the Guttentag Memorial Fund of the Evaluation Research Society.

For information address:

SAGE PUBLICATIONS, INC.
275 South Beverly Drive
Beverly Hills, California 90212



SAGE PUBLICATIONS LTD
28 Banner Street
London EC1Y 8QE

Printed in the United States of America

Library of Congress Cataloging in Publication Data

Main entry under title:
Improving evaluations.

(Sage focus editions; 12)
Includes bibliographies.

1. Evaluation research (Social action programs)—
Congresses. I. Datta, Lois-ellin, 1932- II. Perloff,
Robert. III. Evaluation Research Society. IV. Series.
H62.154 309.1'73'0926 79-13627
ISBN 0-8039-1240-4
ISBN 0-8039-1241-2 pbk.

FIRST PRINTING

CONTENTS

Introduction: Enhanced Lives, Improved Evaluations <i>Lois-ellin Datta and Robert Perloff</i>	9
--	---

SECTION I. OBSTACLES TO EVALUATION

1. Past, Present, and Future Prospects of Evaluation Research <i>Peter H. Rossi</i>	17
2. What Limits the Impact of Evaluations on Federal Policy? <i>Gregory Wurzburg</i>	35
3. Government Constraints on Evaluation Quality <i>Clark C. Abt</i>	43
4. Obstacles to Evaluating State Human Service Programs <i>Michael Simmons</i>	53

SECTION II. IMPROVING EVALUATION DESIGNS

5. Expanding Our View of Control Groups in Evaluations <i>Leonard Saxe and Michelle Fine</i>	61
6. Nonexperimental Evaluation Research: Contributions of Causal Modeling <i>Peter M. Bentler and J. Arthur Woodward</i>	71
7. Some Approaches to Assessing Change <i>Barry S. Tuchfeld</i>	103
8. Applying Time Series Strategies: An Underutilized Solution <i>Michael S. Knapp</i>	111

SECTION III. IMPROVING MEASUREMENT

- | | |
|--|-----|
| 9. Searching for the JND in Mental Health
<i>Charles Windle</i> | 131 |
| 10. Measuring Impact: Power Theory in Social Program
Evaluation
<i>Robert F. Boruch and Hernando Gomez</i> | 139 |
| 11. Against Them or For Me: Community Impact
Evaluations
<i>Harris M. Allen, Jr. and David O. Sears</i> | 171 |
| 12. Cost-Effective Medicaid Mental Health Policies:
Design and Testing
<i>Barbara C. Davenport and Ronald L. Nuttall</i> | 176 |
| 13. Counting Actions Louder Than Words: Behavioral
Scales of Job Performance
<i>Hermann F. Schwind</i> | 185 |
| 14. New Health Technologies: Assessing the Costs and
Benefits
<i>Selma J. Mushkin</i> | 192 |

SECTION IV. THE ROLE OF EVALUATION

- | | |
|---|-----|
| 15. Assumptions Analysis: A Methodology for Strategic
Problem Solving
<i>Ian I. Mitroff, James R. Emshoff, and
Ralph H. Kilmann</i> | 208 |
| 16. Tinkering with Open Systems: Organizational Theory
Perspectives on Educational Program Evaluation
<i>Michael S. Knapp</i> | 226 |
| 17. Emerging Issues for Evaluators and Evaluation Users
<i>Robert F. Rich</i> | 245 |
| 18. Evaluation and Change
<i>Howard Davis and Susan Salasin</i> | 257 |
| About the Authors | 273 |

DEDICATION

*This volume is dedicated
to
MARCIA GUTTENTAG*

*Founding President
Evaluation Research Society*

*A warm and caring colleague
. . . a powerful force for the improvement
and utilization of evaluations as a means toward the
betterment of society and of the human condition.*

INTRODUCTION: ENHANCED LIVES, IMPROVED EVALUATIONS

Lois-ellin Datta and Robert Perloff

In mid-October of 1977, one year after she had presided at the founding of the Evaluation Research Society (ERS), Marcia Guttentag led the opening session of the society's first annual convention. Two weeks later, she suffered a fatal heart attack.

During that first ERS meeting, Marcia was particularly enthusiastic about developing a published series based on the ERS annual meetings.¹ The series—as she conceptualized it—was to be in style between a proceedings, in which the essence of almost every presentation is reported briefly, and an edited book, in which chapters of highest scientific excellence are selected. She hoped for a series which could be used in daily work and in evaluator training to keep evaluation growing and improving.

She particularly wanted to show the connections between evaluation theory and evaluation practice, and among improved theory, improved practice, and the enhancement of human lives. This belief, perhaps more than common endorsement of any technology of evaluation, brings together members of the Evaluation Research Society: a commitment to a just, humane, and humanly rich way of life and the belief

that evaluation can make a contribution to this kind of society worth the energy and commitment of our own lives. This volume² is intended as an experiment in this spirit, and as a memorial to Marcia's own spirit.

The four sections of the volume reflect interests prominent in Marcia Guttentag's published work. The first section, "Obstacles to Evaluation," deals with the political forces influencing what questions are asked, when, and by whom. Rossi begins the section by asserting that evaluation research has fallen into a trap of accepting questions beyond reasonable answering. Citing the negative income tax experiments where political forces beyond the evaluator's control shifted the comparison conditions three times during the course of the five-year experiment, Rossi argues for evaluators' becoming more proactive in identifying the intersect where issues of public interest can be informed by social science research. It is in this intersect that he sees the greatest future contributions of evaluation. Abt looks at the buyer-seller relationship in evaluations, dissecting the limitations unnecessarily imposed by the competitive procurement process in evaluation research. The high cost of submitting proposals relative to the low probability of an award drives out true competition. Once receiving the award, the researcher often is forced to conduct a by-the-recipe study, despite obvious needs for changes in design, measurement, analysis, and reporting from those planned. Abt argues for a change in procurement procedures to better fit the dynamic world in which evaluations take place. Wurzburg suggests that the impact of evaluation results on federal policy is restricted because of certain features of the state of the art of evaluation and because of frequently unstable environments in which policy is created at the federal level. Finally, Simmons offers for consideration fourteen obstacles in the path of evaluating human service programs at the state level.

Writers in the second section, "Improving Evaluation Designs," share a common concern with the question of causal inference or internal validity: how to feel reasonably assured that effects, if any, are attributable to the change under study. The question arises from a Catch-22 quality of many evaluations: in most situations, causal inference is strongest where true controls are possible. In most evaluations, however, true controls are thought to be difficult to obtain, and even when they are possible aberrant control group behavior may make inference more rather than less difficult. In her many struggles with this problem,

Guttentag was led to inverse probability as a mathematical model which best fit the political and philosophical beliefs she cherished and her view of what was realistic: a shifting, changing reality where those most directly affected by the evaluation findings should have greatest control over what was studied, and would have their values most prominently reflected in the results. While none of the articles in this section springs from the Multi-Attribute Utility Theory or Decision-Theoretic approach which she and her colleagues developed, the definition of the problem and the search for alternative solutions were topics she found exciting and important.

Saxe and Fine begin with a model for evaluation which embeds control group designs for answering specific questions within a larger evaluation model. They suggest that the either/or question of a control group versus the decision-theoretic approach is artificial and unproductive. Their case study illustrates one application of a combined model, an attempt to gain the strongest inference possible about impact in a study of television diagnosis in rural areas, while retaining the pragmatic convenience of the nonexperimental models for answering other questions. The three other papers in this section range from the true experimental design to statistical paradigms for drawing inferences. Bentler and Woodward present an approach to causal modeling in which process data are used to select the most plausible of alternative interpretations of changes associated with an intervention. Knapp describes another strategy: the use of time series data to rule out competing hypotheses and establish causality. Like Bentler and Woodward, Knapp identifies the limiting conditions of this approach, while arguing that it is a presently an underutilized and extremely powerful method. Tuchfeld, closing this section, offers an overview of statistical methods for measuring change. His survey of these methods presents several facets of the dilemma (faced squarely by Guttentag) between design and analysis in drawing causal inference.

"Improving Measurement," the third section of the volume, deals with three concerns. Windle's paper raises the question of how large an effect is needed in order to be practically, as well as statistically, meaningful. Although his examples are focused on mental health program evaluations, the often-neglected distinction between statistical significance (which is fairly easy to change by increasing N s and the reliability of the measures) and practical significance could be applied

to other fields. Boruch and Gomez treat a second concern: the impact of measurement unreliability on the power of evaluations to detect differences when they actually exist, and what statistical approaches can be applied to gain power. The third issue is the construction and interpretation of measures in different fields. Allen and Sears, in a case study of desegregation, illustrate the care needed in interpreting attitudinal data. Schwind, in the context of bank manager training, presents a psychometric improvement on the use of task sampling for scale construction. Davenport and Nuttall, dealing with mental health programs supported by Medicaid, suggest ways of thinking about values in a field where a positive outcome may be an increase in clients rather than the more usual decreases in number of cases and time spent on each. Mushkin, facing a similar dilemma, illuminates hitherto murky areas of thinking about costs and benefits of the new, life-sustaining health technologies.

These areas, too, Marcia found exciting. Her own work at first glance more frequently dealt with design and analysis than with improving measures per se. Looked at in more detail, almost all her studies grappled with trying to think clearly about what the numbers meant: about conceptually establishing important social issues involved in *any* of the content areas; about creating imaginative, sensitive approaches to studying these conceptually salient questions; and about great care in translating back from the computer printouts to meaningful implications for the people and programs reflected, however dimly, in the quantitative data. In her own work, conceptual sense and creative assessment were more dominant than psychometrics for its own sake. While recognizing the need for valid, reliable measures, she would place higher value on meaningful, interpretable data, a concern of the papers in this section.

In "The Role of Evaluation," section four, the papers return to issues raised in the first section, to the relation of evaluation to the social order in which it is inextricably embedded. While the first section deals with government-imposed obstacles to evaluation, the fourth section turns to what evaluators may uniquely contribute to improving the impact of data on policy and practice. Mitroff, Emshoff, and Kilmann consider the role of assumptions in evaluations: the preference of many evaluators for specificity in defining the program objectives and purposes despite the multi-purpose, multi-actor nature of even the smallest

program with the most narrowly defined purpose. They present a heuristic for dealing with this, which involves a systematic testing of strategic assumptions. Knapp, in his examination of organizational theory, offers an approach more attuned to human relations. He argues that most social systems in education pull toward the individual and away from the group. Evaluators, rather than seeking group consensus, he argues, should identify the opportunities for changes and design evaluations to reflect these concerns. Rich identifies a variety of critical issues that evaluators need to address, including the availability and utilization of appropriate methodological tools, the identification of training programs suitable for preparing evaluators to face the realities and challenges in evaluation tasks as scientists and practitioners, the continuing need to translate evaluation results into humane and relevant policy, the need to reduce instances of abuse and misutilization of evaluation data, and the development of ethical standards for evaluators. Finally, Davis and Salasin step back from these specific issues in their analysis of the relation between evaluation and change. They argue primarily that evaluators should accept a role as change agents. Rather than lamenting that evaluations are not utilized (when they are dropped from Olympus), evaluators should come down to earth, using presently available techniques for understanding readiness for change, and better meshing the evaluation studies with the situations as they find them—in their human and humane diversity, inconsistencies, and complexities.

In some ways, these chapters probably would have worried Marcia. She cared deeply about change, and spared neither herself nor others in the directness of her efforts to see justice done and individuals enabled to create their own lifestyle in their own life space. She hoped evaluation, by giving individuals information about themselves, their situation, and about approaches that worked, could be, if necessary, a terrible swift sword, but preferably a plowshare. She would have resisted whatever would have compromised with might against right, or placed evaluation in the position of background musician among the counselors of the princes. Yet in many other ways, these efforts, like her own, to find ways in which evaluations could be most useful in promoting humane, moral, life-enhancing change—both within the profession and in society—were like her own, and might well have earned her brightest smile and her heartiest “Good work.”

NOTES

1. As this volume goes to press, we see the realization of Marcia Guttentag's aspirations in the establishment of the forthcoming (commencing in fall 1979) Sage Research Progress Series in Evaluation—a joint undertaking of the Evaluation Research Society and Sage Publications, under the able editorship of Susan Salasin. The series will consist of brief, theme-oriented volumes (developed in part from papers presented at each year's ERS meetings) for the use of evaluation researchers, practitioners, and students.

2. *The essays in this volume were selected not only for their quality but also for their relation to the volume's major themes. We are grateful to the contributors—both those who wrote new essays for this volume and those who cooperated so cheerfully with us in the revision process. The volume's royalties will be donated to the Society's Guttentag Memorial Fund to further advance us toward the goals which Marcia Guttentag sought to reach.*

SECTION I
OBSTACLES TO EVALUATION

