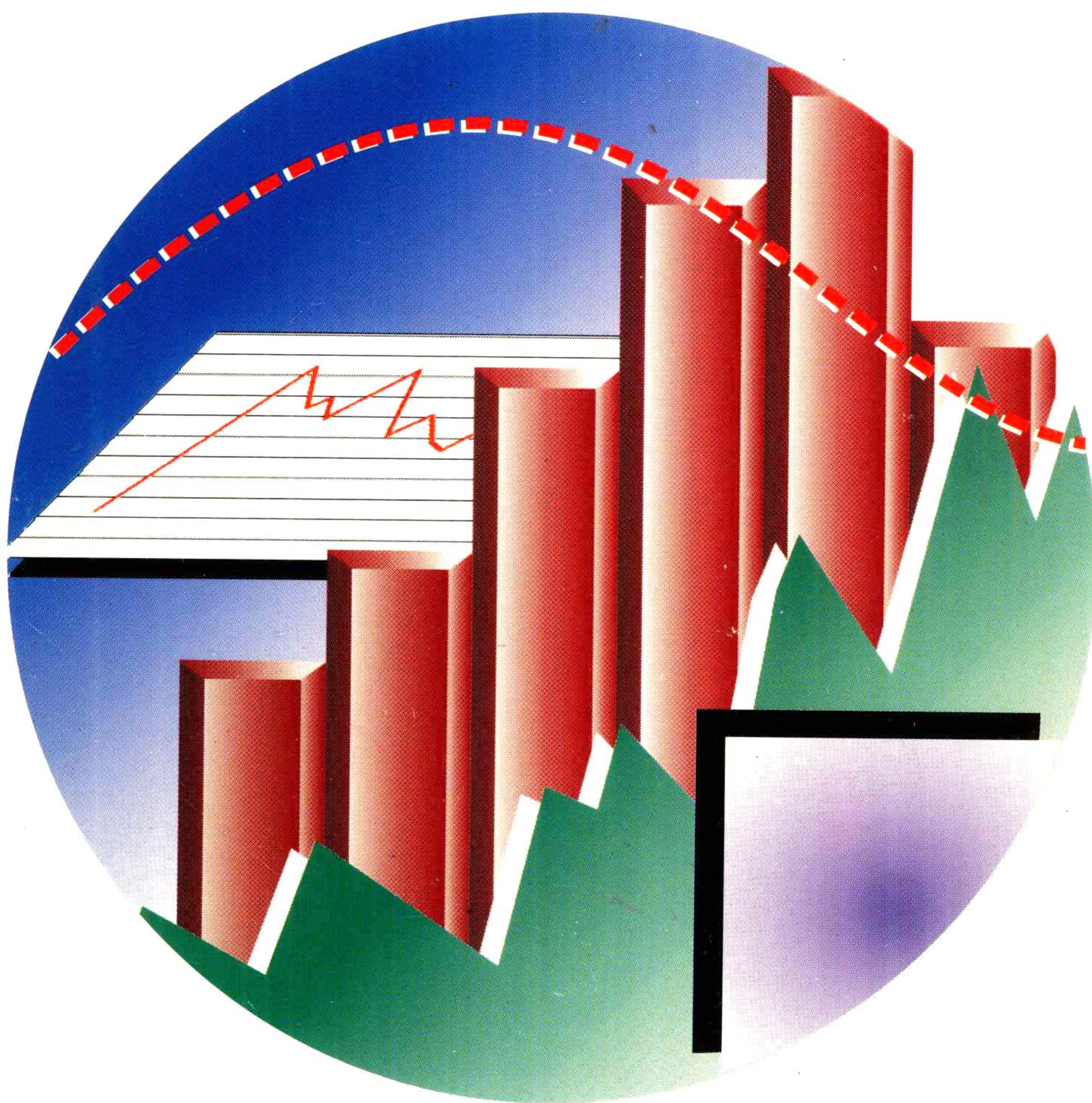


Health Education Evaluation and Measurement

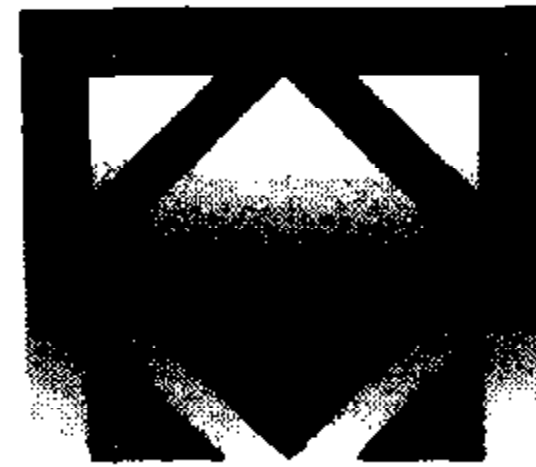
*A Practitioner's
Perspective*



Paul D. Sarvela ■ Robert J. McDermott

Health Education Evaluation and Measurement

A Practitioner's Perspective



Paul D. Sarvela
Southern Illinois University at Carbondale

Robert J. McDermott
University of South Florida



Boston, Massachusetts Burr Ridge, Illinois Dubuque, Iowa
Madison, Wisconsin New York, New York San Francisco, California St. Louis, Missouri

Book Team

Editor *Chris Rogers*
Developmental Editor *Susan J. McCormick*
Production Coordinator *Peggy Selle*

WCB/McGraw-Hill

A Division of The McGraw-Hill Companies

Vice President and General Manager *Thomas E. Doran*
Executive Managing Editor *Ed Bartell*
Executive Editor *Edgar J. Laube*
Director of Marketing *Kathy Law Laube*
National Sales Manager *Eric Ziegler*
Marketing Manager *Pamela Cooper*
Advertising Manager *Jodi Rymer*
Managing Editor, Production *Colleen A. Yonda*
Manager of Visuals and Design *Faye M. Schilling*

Production Editorial Manager *Vickie Putman Caugbron*
Publishing Services Manager *Karen J. Slaght*
Permissions/Records Manager *Connie Allendorf*

Chairman Emeritus *Wm. C. Brown*
Chairman and Chief Executive Officer *Mark C. Falb*
President and Chief Operating Officer *G. Franklin Lewis*
Corporate Vice President, Operations *Beverly Kolz*
Corporate Vice President, President of WCB Manufacturing *Roger Meyer*

Cover and interior design by John Rokusek

Copyedited by Julie Bach

Copyright © 1993 by Wm. C. Brown Communications, Inc. All rights reserved

Library of Congress Catalog Card Number: 91-76691

ISBN 0-697-12769-9

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

Printed in the United States of America

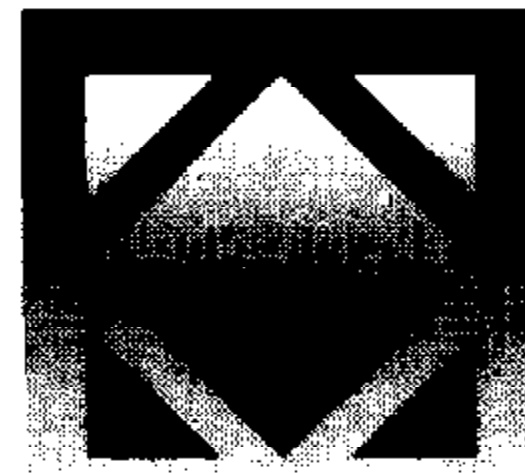
567890 QPD 0987

Health Education Evaluation and Measurement

A Practitioner's Perspective

*For Debbie, Kristin, and Kay
and in memory of Mac (1921-1991)*

Foreword



Evaluation is a process that each of us engages in daily. We examine, critique, and make choices on an on-going basis. This kind of evaluation, however, is not what comes to mind when thinking about health education or health promotion interventions. To conduct this type of evaluation, we must design a strategy to answer questions of importance and use that information to contribute to the knowledge base. Evaluation as a systematic process provides a method to answer critical questions. Although there are many different strategies for conducting evaluation activities in health education, there are some commonalties found among them, including identifying appropriate questions to be answered; formulating a plan designed to provide answers to these questions; gathering appropriate data according to that plan; examining these data for their meaning; interpreting that meaning in direct relationship to the questions posed initially; and, reporting the results in a manner that maximizes the potential for utilization of results.

Although many different strategies may be used in evaluation activities, the six steps discussed above are found in almost all instances. This framework, however, does not eliminate the potential problems inherent in these strategies. The current literature identifies numerous categories related to these problems: 1) issues in measurement and instrument design, including reliability and validity concerns; 2) lack of understanding of utilization issues; 3) ethical issues in research, including fraud and deception, freedom of information, confidentiality, anonymity, and clinical trials.

Overcoming these and other problems is the subject of this book by Paul Sarvela and Robert McDermott. Although there are several books that someone interested in evaluation of health education/health promotion interventions may choose, this book represents a unique addition to the literature base in four ways:

1. ***It is written for the real world.*** Most introductory evaluation books are designed to provide an overview of how a practitioner can conduct evaluation efforts. It has been my belief that evaluation activities, as with research activities, represent a compromise

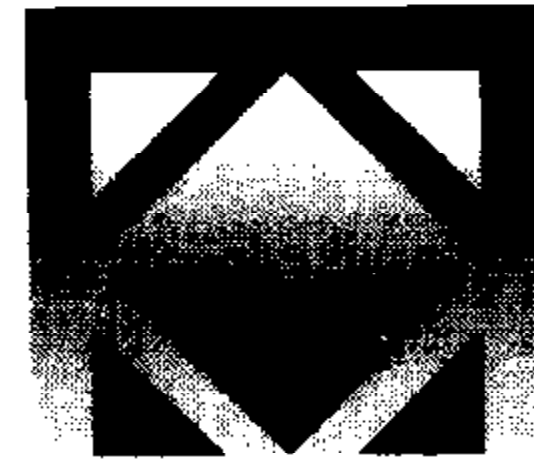
between the ideal and the reality of working in environments that cannot be controlled. Therefore, research design should be defined as a series of systematic compromises between what is desirable and what is realistic in any given research situation. Most evaluation books talk more about the ideal than what reality allows.

2. *It is written in a "how to" format.* One of the most important features of this book is its clear definitions and the listing of steps required to complete the activities necessary for evaluation.
3. *It is written for the "front-line" practitioner.* Most evaluation books appear to be written for the person who is charged with evaluating large, well-funded projects. Most evaluation books make the assumption that the reader is already skilled in many of these areas. This book assumes that the person using this book is a front-line practitioner who needs to know how well a project is going and how to make it better.
4. *It contains chapters not found in other books or chapters with a different emphasis than is traditional.* The chapter on qualitative methods adds a substantive dimension to this book missing in many others. In addition, rather than wait until the end of a book to discuss the ethical issues as an add-on, Sarvela and McDermott correctly recognize their importance in the planning, design, and implementation of evaluation activities. As a result, it becomes critical to review these concerns at the beginning of a discussion on evaluation.

A major goal of the book, as described by the authors, is to make it "user friendly." While others may define this term differently, it is my belief that a book written for the real world in a how-to format, for front-line practitioners, and with new and innovative approaches to information is a fine operational definition of "user-friendly".

Bob Gold
University of Maryland
June 1992

Preface



Our task was to write a “user-friendly” text. We tried to write a text that would be useful to three groups of people: students, teachers, and health educators working in the so-called “real world.”

We have found that most of the students in an evaluation and measurement class are there because they are required to be there. Through our conversations with colleagues in academia, we have discovered that those who teach evaluation are often assigned the class by their supervisors, an assignment often accepted with some reluctance. In addition, the practitioners of health education are often “asked” by their supervisors to conduct an evaluation. As with teachers, practitioners may not volunteer eagerly to be in charge of an evaluation project.

Many people who are asked to study, teach, or conduct program evaluation are not happy to do so because the reference texts and research reports are impenetrable, due to their emphasis on psychometrvia and esoteric statistical procedures. We readily acknowledge that many of the current evaluation and measurement textbooks are technically sound, describing in detail how to conduct large-scale community intervention evaluations. However, we also believe that most people will *never* conduct a large-scale evaluation project. The bulk of their evaluation activity will occur on a much smaller scale, dealing with pilot-testing a program or conducting a satisfaction survey. We have written primarily to this audience.

Evaluation is an exciting topic and should be taught in an exciting manner. Pick up any newspaper, listen to the radio, or watch TV and chances are great that you will hear the results of an evaluation study dealing with health, education, the environment, or the economy. Evaluation is a topic that is alive, important, and stimulating. In that spirit, we have provided background material, case studies, and references of reports and studies on interesting and controversial issues. These resources are used to illustrate that we are all affected by evaluation studies. At the end of each chapter, we have provided a series of questions and activities. These exercises are meant to stimulate thought and discussion concerning important evaluation

questions facing health education and health promotion specialists today. By presenting the basic information concerning evaluation and measurement, and then asking the reader to think about and discuss important evaluation and measurement issues, we feel that the relevance of the topics covered in the text will be understood better.

The text is comprised of fourteen chapters. The introductory chapter provides a brief review of the purposes and history of evaluation along with a discussion on the objectives of evaluation. Chapter two focuses on the politics and ethics of evaluation. Evaluation is a politically explosive process, laden with ethical concerns and issues such as the evaluator-client contract and responsibilities of the evaluator, and balancing the needs of different stakeholders. Chapters three, four, and five address measurement, focusing on types of instruments used in health education evaluation, steps in the development of an instrument, reliability and validity, and methods for measuring knowledge, attitudes, and behavior. Quantitative designs, covering nonexperimental, quasiexperimental, and true experimental methods, along with threats to internal and external validity, are discussed in chapter six. Chapter seven focuses on qualitative methods with a special emphasis on commonly used qualitative techniques.

We felt that pilot-testing was so important that it warranted its own chapter, so chapter eight focuses on the purposes and methods used for pilot testing instruments, data collection and analysis procedures, and curriculum materials. Chapter nine examines the topic of needs assessment and strategic planning, providing an overview of commonly used needs assessment and planning models as well as possible sources for data. An examination of program costs in relation to program effects and benefits is becoming an increasingly common effort for evaluation specialists. These issues are discussed in chapter ten. A discussion of different types of sampling designs and the selection of sample size is found in chapter eleven. The logistics of evaluation, an often overlooked but extremely important part of planning for evaluations, is covered in chapter twelve, while a readable examination concerning statistical analysis of data is provided in chapter thirteen. The text concludes with a chapter on reporting and using evaluation data.

You will note from the above review of the chapters that we have included standard chapters on measurement and quantitative and qualitative methods, but also have highlighted a few other areas not often addressed thoroughly in current evaluation and measurement texts, such as the politics of evaluation, pilot-testing, needs assessment and strategic planning, and the logistics of evaluation. It is our hope that the addition of these chapters will help practitioners carry out evaluations in a more effective manner.

We tried to write this text with the reader in mind. Therefore, each chapter begins with an introduction, a list of chapter objectives, and key terms to serve as an “advanced organizer” for the material that follows. To stimulate further thought on the topics as well as provide closure, each chapter ends with a case study and set of questions or activities. The case studies, questions, and activities can be considered individually by the reader, or of course, in a classroom setting, assigned to small groups as a classroom activity.

We enjoyed writing this book. In fact, we are still good friends at the completion of this text (it has been said that there are three ways to lose a good friend: have him date your sister, sell him your car, or write a book with him). If you enjoy reading this text, and learn a little about evaluation and measurement in the process, we have met our goal.

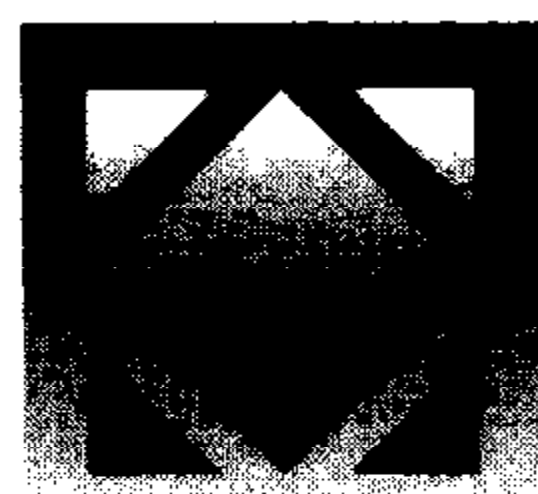
Acknowledgements

We wish to thank the following colleagues for their help in the preparation of this manuscript: Julie Doidge Huetteman, Robert Weiler, Mark Kittleson, Derek Holcomb, and Carol Bryant. We also would like to thank James Eddy, University of Alabama; Cathy Kennedy, Colorado State University; Mary Sutherland, Florida State University; Mohammed Torabi, Indiana University; and David White, East Carolina University, who served as outside reviewers for the text, and Chris Rogers, Susie McCormick, Peggy Selle, and Rosemary Wallner for their fine editorial efforts.

Paul D. Sarvela, Carbondale, IL

Robert J. McDermott, Tampa, FL

Contents



Foreword xi

Preface xiii

1 Introduction to Program Evaluation 1

Chapter Objectives 1

Key Terms 1

Introduction 2

A Historical Overview of Program
Evaluation 3

History of Health Education
Evaluation 6

The Purposes of Evaluation 6

Formative and Summative
Evaluation 10

The Foci of Evaluation 11

Summary 12

Case Study 13

Student Questions/Activities 13

References 14

2 The Politics and Ethics of Program Evaluation 15

Chapter Objectives 15

Key Terms 15

Introduction 16

Evaluation: Who Should Do the
Job? 16

The Evaluator-Client Contract 19

The Basic Needs of Stakeholders 20

The Political Nature of Evaluation 22

The Ethics Continuum in Program
Evaluation 27

The Responsibilities of Institutional
Review Boards 30

Summary 31

Case Study 32

Student Questions/Activities 32

References 33

3 Introduction to Measurement 35

Chapter Objectives 35
Key Terms 35
Introduction 36
Contemporary Problems in Measurement 36
Levels of Measurement 38
Models, Frameworks, Curriculum Materials, and Instrument Development 40
Types of Measurement Instruments Used in Health Education 42
Steps in the Development of an Instrument 44
Summary 50
Case Study 51
Student Questions/Activities 51
References 52

4 Reliability and Validity 54

Chapter Objectives 54
Key Terms 54
Introduction 55
Reliability 55
Validity 64
Methods of Enhancing Reliability and Validity 70
Summary 77
Case Study 77
Student Questions/Activities 78
References 79
Appendix 4.1 81

5 Types of Measures 83

Chapter Objectives 83
Key Terms 83
Introduction 84
Measuring Knowledge 84
Measuring Attitudes 90

Measuring Behavior 97
Common Medical Tests and Procedures 100
Setting Cut Scores 102
Computer-Assisted Testing 104
Summary 106
Case Study 106
Student Questions/Activities 106
References 106

6 Quantitative Evaluation: Models, Methods, and Designs 109

Chapter Objectives 109
Key Terms 109
Introduction 110
Theory of Quantitative Evaluation 110
Internal and External Validity 111
Threats to Internal Validity 111
Threats to External Validity 117
Non-Experimental Evaluation Methods and Designs 119
Experimental and Quasi-Experimental Methods and Designs 122
Summary 129
Case Study 129
Student Questions/Activities 129
References 130

7 Qualitative Evaluation: Models, Methods, and Designs 132

Chapter Objectives 132
Key Terms 132
Introduction 133
Qualitative Evaluation in Perspective 133
Commonly Used Designs and Models 135
Qualitative Methods 139

Summary 150
Case Study 150
Student Questions/Activities 151
References 151

8 Pilot-Tests 155

Chapter Objectives 155
Key Terms 155
Introduction 156
The Purposes and Levels of Pilot-
Tests 156
Instruments 160
Data Collection and Sampling
Procedures 163
Data Analysis Procedures 165
Curriculum Materials and
Instructional Strategies 166
Project Cost 174
The Revision Process 174
Summary 175
Case Study 175
Student Questions/Activities 176
References 176
Appendix 8.1 177

9 Needs Assessment and Strategic Planning 184

Chapter Objectives 184
Key Terms 184
Introduction 185
The Purposes of Needs Assessment
and Strategic Planning 187
Definition of Need 188
Needs Assessment and Strategic
Planning Models 190
Data Sources and Methods for Needs
Assessments 197
Summary 199
Case Study 199
Student Questions/Activities 199
References 199

10 Cost Analysis 201

Chapter Objectives 201
Key Terms 201
Introduction 202
Models of Cost Analysis 204
Strengths and Weaknesses of Cost
Analysis 214
Summary 217
Case Study 217
Student Questions/Activities 217
References 218

11 Methods and Strategies for Sampling 220

Chapter Objectives 220
Key Terms 220
Introduction 221
What Is Sampling? 221
Nonprobability Samples 222
Probability Samples 228
Additional Methods of Probability
Sampling 232
Nonresponse: Then What? 238
Sample Size: How Many Subjects Do I
Need? 239
Generalizing from Samples 240
Summary 241
Case Study 241
Student Questions/Activities 242
References 242

12 The Logistics of Evaluation 244

Chapter Objectives 244
Key Terms 244
Introduction 245
Planning and Managing the
Evaluation 245

The Contract/Professional Agreement 247
 Other Contractual Considerations 250
 Budget Preparation 252
 Logistics of Data Collection 257
 Planning Aides 259
 Use of Computers in Evaluation 262
 Summary 271
 Case Study 272
 Student Questions/Activities 272
 References 274

13 Data Analysis 275

Chapter Objectives 275
 Key Terms 275
 Introduction 276
 Univariate Procedures 277
 Bivariate Procedures 283
 Multivariate Procedures 288
 Vital Statistics 290
 Summary 292
 Case Study 293
 Student Questions/Activities 293
 References 293
 Appendix 13.1 295

14 Preparing the Evaluation Report 296

Chapter Objectives 296
 Key Terms 296
 Introduction 297
 Content of the Evaluation Report 297
 Front Matter 298
 The Executive Summary 300
 Program Background and Evaluation Description 301
 Presentation of Results 302
 Discussion, Conclusions, and Recommendations 307
 Summary 310
 Case Study 311
 Student Questions/Activities 314
 References 314

Glossary 315

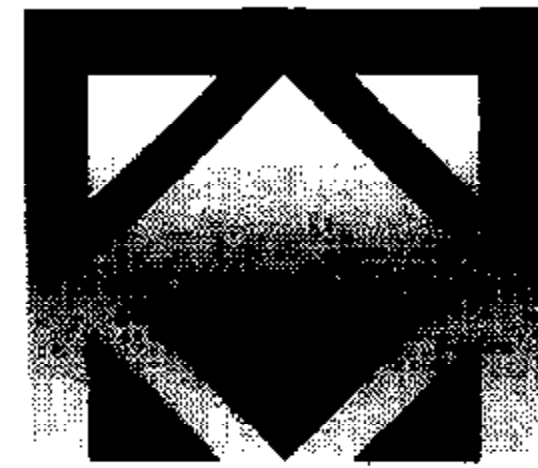
Name Index 323

Subject Index 329

Chapter

1

Introduction to Program Evaluation



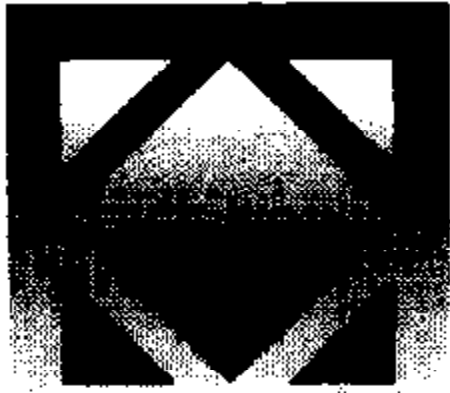
Chapter Objectives

*After completing this chapter, the reader
should be able to:*

1. Define evaluation from several perspectives.
2. Identify key events in the history of evaluation.
3. Describe the purposes of evaluation.
4. Distinguish among the types of evaluation.
5. Identify the foci of evaluation.
6. Explain why evaluation of health education programs is both timely and important.

Key Terms

formative evaluation
impact evaluation
outcome evaluation
process evaluation
program evaluation
quality assurance
quality control
summative evaluation



Introduction

Benjamin Franklin once wrote:

"In this world, nothing can be said to be certain except death and taxes."

Today, there should be a modification to this statement:

"In this world, nothing can be said to be certain except death, taxes, and being evaluated."

From the moment we are born to the moment we die, we are evaluated.

At birth a baby is given a number of tests. The baby is measured and weighed, checked for birth defects, and given an Apgar test at one and five minutes that measures basic physiological functioning. The physician uses these and other data to determine the course of treatment for the newborn. The data also are reported on the birth certificate, which is sent to the local health department. Therefore, not only is the baby being evaluated, but so is the general public health of the community. The birth certificate data are analyzed to determine prevalence of such things as premature births, low birth weight babies, birth defects, and demographic characteristics of the parents. Extending the discussion a step further, one may argue that today, babies are frequently evaluated even before they are born, using sophisticated procedures such as ultrasound, amniocentesis, and a host of biomedical tests that can be performed to determine the health status of the baby.

The cycle of evaluation continues throughout life. We are tested (some would say over-tested) throughout our school years. Again, not only is the individual evaluated when taking an important test (e.g., when taking an ACT or SAT test), but so are the schools. Every year, newspapers report how local school districts compare to each other in terms of standardized test scores, and how the local schools compare to state and national averages.

One might finish university studies, fraught with evaluation throughout the course of study, and then be required to take a licensing or certification test (e.g., an exam to be a licensed professional engineer, a registered nurse, or a certified health education specialist) in order to work in the profession for which one was prepared. Persons who join the armed forces face a barrage of tests (e.g., medical physicals and fitness, aptitude, and qualification tests). As with the earlier testing, not only is the individual being evaluated, but so is the organization with which the individual is associated. Nursing programs, law schools, medical schools, and other professional schools are frequently evaluated, in part, by how many of their students pass licensing and certification examinations.

Even during the final stages of life, and thereafter, one is evaluated. The physician tests our brain to see if we are dead. Once we die, our age, cause of death, and other personal characteristics are reported on a death certificate. And of

course, these data are reported to the health department, so, again, our community can be evaluated in terms of mortality rates as compared to the norms established by state and federal health agencies.

Clearly, evaluation is an important part of our daily life. Evaluation is found in some form in all professions. In business and industry it is known as quality control or quality assurance. Many school systems employ evaluators in some capacity, as public expectations concerning the accountability for educational programs increases. We continuously evaluate medical care in terms of its effectiveness. Like other professionals, health educators view evaluation as an important part of delivering high quality programs to their target populations.

Just as health education evaluators work in many different settings, they also are charged with many different tasks. The achievement of each of these tasks can be evaluated in different ways. Through this text, we seek to introduce you to some basic elements of conducting effective and defensible evaluations of health education and promotion programs and materials.

A Historical Overview of Program Evaluation

Program evaluation is often viewed as a recent phenomenon, first achieving prominence in the 1960s. However, program evaluation has taken an active role in the educational process for thousands of years. S. M. Shortell and W. C. Richardson (1978) indicate that the evaluation of medical programs has taken place for centuries. Today's evaluations are related to early sanctions in Egypt in 3000 B.C. If a patient lost his or her eye unnecessarily, the physician could lose a hand. B. R. Worthen and J. R. Sanders (1987) note that Chinese officials, as early as 2000 B.C., used civil service examinations to measure proficiency of public officials. They further indicate that Socrates and other early Greeks used verbal exams as an important part of their teaching methods.

G. F. Madaus, D. L. Stufflebeam, and M. S. Scriven (1983) describe the evolution of evaluation in six stages:

1. The Age of Reform (pre-1900)
2. The Age of Efficiency & Testing (1900-1930)
3. The Tylerian Age (1931-1945)
4. The Age of Innocence (1946-1957)
5. The Age of Expansion (1958-1972)
6. The Age of Professionalization (1973-present)

The Age of Reform took place before the 1900s. Great Britain was probably one of the first countries to be involved formally in program evaluation. In 1870, the Royal Commission of Inquiry into Primary Education in Ireland, after conducting an evaluation based on testimony and examining evidence, indicated:

"The progress of the children in the national schools of Ireland is very much less than it ought to be."