

JOHN BOULMETIS

PHYLLIS DUTWIN

THIRD EDITION

THE ABCs OF
EVALUATION

TIMELESS TECHNIQUES FOR

PROGRAM AND PROJECT MANAGERS

The ABCs of Evaluation

Timeless Techniques for Program and
Project Managers

Third
Edition

John Boulmetis
Phyllis Dutwin

 **JOSSEY-BASS**
A Wiley Imprint
www.josseybass.com

Copyright © 2011 by John Wiley & Sons, Inc. All rights reserved.

Published by Jossey-Bass

A Wiley Imprint

989 Market Street, San Francisco, CA 94103-1741—www.josseybass.com

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, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400, fax 978-646-8600, or on the Web at www.copyright.com. Requests to the publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, 201-748-6011, fax 201-748-6008, or online at www.wiley.com/go/permissions.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Readers should be aware that Internet Web sites offered as citations and/or sources for further information may have changed or disappeared between the time this was written and when it is read.

Jossey-Bass books and products are available through most bookstores. To contact Jossey-Bass directly call our Customer Care Department within the U.S. at 800-956-7739, outside the U.S. at 317-572-3986, or fax 317-572-4002.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Not all content that is available in standard print versions of this book may appear or be packaged in all book formats. If you have purchased a version of this book that did not include media that is referenced by or accompanies a standard print version, you may request this media by visiting <http://booksupport.wiley.com>. For more information about Wiley products, visit us at www.wiley.com.

Library of Congress Cataloging-in-Publication Data

Boulmetis, John.

The ABCs of evaluation : timeless techniques for program and project managers / John Boulmetis, Phyllis Dutwin.—3rd ed.

p. cm.—(Research methods for the social sciences ; 46)

Includes bibliographical references and index.

ISBN 978-0-470-87354-0 (pbk.); 978-1-118-10364-7 (ebk.); 978-1-118-10365-4 (ebk.); 978-1-118-10366-1 (ebk.)

I. Management—Evaluation. I. Dutwin, Phyllis. II. Title.

HD31.B633 2011

658.4¹032—dc23

2011021314

Printed in the United States of America

THIRD EDITION

PB Printing 10 9 8 7 6 5 4 3 2 1

CONTENTS

Preface vii

The Authors xv

1 What is Evaluation? 1

- Common Denominators 3
- Two Definitions of Evaluation 4
- Evaluating Efficiency, Effectiveness, and Impact 5
- Evaluating Alternatives 8
- Identifying Areas to Improve 8
- Other Levels 10
- Formal Reasons to Evaluate 12
- Practical Application 15
- Formative and Summative Evaluations 16
- Putting It All Together 16
- Key Words and Concepts 24
- Questions and Exercises 24
- Further Reading 27

2 Why Evaluate? 29

- Grants and Grantsmanship 31
- Benefits 33
- Limitations 35
- Considerations 36
- Putting It All Together 40

Key Words and Concepts 40
Questions and Exercise 43
Further Reading 46

3 Decision Making: Whom to Involve, How, and Why? 47

Program and Project Cycles 49
Evaluation and Monitoring 52
Program Planning Cycle and Evaluation 57
Putting It All Together 61
Key Words and Concepts 65
Questions and Exercises 65
Further Reading 73

4 Starting Point: The Evaluator's Program Description 75

What an Evaluator's Program Description Can Do 77
Developing an Evaluator's Program Description 78
Putting It All Together 83
Key Words and Concepts 84
Questions and Exercises 84
Further Reading 96

5 Choosing an Evaluation Model 97

More Than One Way to Evaluate 98
Four Evaluation Models 101
Other Models 108
Choosing the Right Model 110
Evaluation Design Format 111
Other Considerations 114
Putting It All Together 114
Key Words and Concepts 115
Questions and Exercise 115
Further Reading 117

6 Data Sources 119

Starting Point 120
Qualitative and Quantitative Data Sources 124
Four Levels of Data 125
Nature of Data Collection 127
Existing Data 128
Newly Collected Data 130

Putting It All Together 141
Key Words and Concepts 141
Questions and Exercise 141
Further Reading 144

7 Data Analysis 147

Data Analysis 148
Data 150
Terms 151
Levels of Data 157
Putting It All Together 164
Key Words and Concepts 164
Questions and Exercise 165
Further Reading 167
Software for Personal Computers 168

8 Is It Evaluation or Is It Research? 169

Determining Evaluation or Research from Intent 170
Determining Evaluation or Research from Design 172
Qualitative and Quantitative Characteristics 174
Sampling 175
Sampling and Surveys 180
Putting It All Together 181
Key Words and Concepts 183
Questions and Exercise 183
Further Reading 184

9 Writing the Evaluation Report 185

Communicating Your Findings 187
Evaluation Reports 188
The Report Outline 192
Putting It All Together 201
Key Words and Concepts 202
Questions and Exercise 202
Further Reading 203

10 Evaluation as a Business 205

Is There a Market for Program Evaluation? 206
How Do You Find the Business? 211
Proactive Resources 213

Reactive Resources	216
Emergency Resources	216
How Do You Negotiate a Contract?	217
Keeping the Business Alive	222
Putting It All Together	224
Key Words and Concepts	225
Further Reading	225

A: More on Data Analysis 227

Analyses for Each Level	227
Significance Testing	231
Statistical and Practical Significance	235

B: Sample Evaluation Report 237

C: Sample Evaluation Contract 271

References 277

Index 281

PREFACE

Evaluation is a task that every program and project manager will face at one time or another: you may manage the funding process for an organization, or direct a self-study that requires evaluation, or write a grant proposal that includes an evaluation segment. Occasions vary, but they are inevitable.

Still, endorsing evaluation is a lot like endorsing regular visits to the dentist. People are quick to endorse both activities, but when it comes to doing either one, most people are very uncomfortable. *The ABCs of Evaluation* will reduce your discomfort by doing two things: First, the book will demystify the process of designing and conducting an evaluation by helping you understand the components of an evaluation design. Second, and more important, the book's aim is to convince you that you play an integral part in any evaluation process, and therefore you are a beneficiary of the results.

Who is the *you* that this book is designed to assist in the process of evaluation? You might be a manager, supervisor, team facilitator, analyst, or evaluator in the private sector. Then again, you might be in the public or private sector as an administrator, coordinator, facilitator, consultant, or evaluator. Or you might be a faculty member or student in any of a host of disciplines (education, management, human services,

psychology, public affairs, labor relations, or health). For any of these situations, the book will provide you with a basic understanding of the steps to follow to design, conduct, and supervise a program evaluation.

This volume has been designed to serve as an introduction to evaluation for program and project managers who have little experience in this task. It presents time-tested evaluation principles and techniques, but it does not purport to reflect the totality of the professional field of evaluation, which encompasses a far greater range of techniques and technologies than are suited to the needs and interests of this book's readership.

Demonstrating Success

As a professional, you know that your programs need to show evidence of success. For example, regardless of how important a communication skills training seminar might be to the overall team-building effort of a company, if you cannot show how it affected the employees' teamwork skills or how improved teamwork added to company performance, the program might not be funded again. This realization may have originated with you, or the organization may require evaluations and demand accountability. If you have always assumed that evaluations are just too difficult for you to do, that you don't want to tackle difficult data collection or unfathomable data analysis, think again. *The ABCs of Evaluation* will dispel these misconceptions and show you what to do.

Both the person who is designing and performing an evaluation and the person who is participating in and receiving the findings of an evaluation need to prepare themselves to understand the basic processes involved. The evaluator, by understanding the thoughts, concerns, and questions of the evaluated, can better anticipate and prepare to address them up front. The evaluated, by anticipating the processes, the probing, and the data requests of an evaluator, can position themselves to make better use of the evaluation process and results.

Too often people feel that evaluation is something that is done to them rather than for them. Indeed, evaluation may be seen as something to get through and get around (Gray, 1998).

The ABCs of Evaluation is an easy-to-read explanation of the concepts and methods of the evaluation process. The chapters are designed to teach readers what the process offers, what it consists of, and how it is designed and implemented. In addition, the book will guide managers step-by-step throughout the process of conducting an evaluation, from the early question, “Why evaluate?” to the later one, “What did the evaluation tell me?” Consequently, the book will be useful to those who design and conduct the evaluation, as well as to those who supervise others in these tasks.

You may be experiencing downsizing, reorganization, or the introduction of high-performance work teams where you work. These new circumstances may require you to possess evaluation skills. Although you now find yourself in a position that requires the skills, you may not have them, or you may not have been trained to use them.

Overview

At the beginning of each chapter, a scenario sets the stage. Then the chapter presents new evaluation concepts followed by concluding exercises designed to challenge your understanding.

Chapter One defines evaluation, keeping in mind that it is considered both an art and a science. The chapter proposes that readers learn principles and theories as a first step and then apply them in actual situations.

In Chapter One, you will learn that evaluation, both quantitative and qualitative, means measuring and collecting data against some standard. The evaluation may be of people, any activity, an entire program, or all of the above.

After an introduction to the formal reasons to evaluate (more on this in Chapter Two), the chapter explores the ideas of setting standards, using project cycles, and determining levels in evaluation.

Finally, there is an introduction to the evaluation design format that the reader will learn to use throughout the book.

Chapter Two encourages you to think about the following: where there is a program, there must be evaluation. The “Why?” of evaluation emanates from different sources or audiences—from a need you have, or from a requirement that administration or management has for planning, policymaking, funding, or ongoing research. Whatever its purposes and objectives, evaluation is an integral part of most programs and should be planned from the outset. The evaluation design format introduced in Chapter One comes into use here to address your need to define your audience and to eventually report to that audience.

Chapter Three explores the importance of management buy-in to decision making and explores how staff, subject-matter experts, and other stakeholders are involved in the preplanning, planning, and application of the evaluation. The chapter also addresses this question: How can I evaluate a program I did not help create?

Evaluation decisions flow through the various stages in the evaluation. Consequently, in this chapter you will learn whom to involve in the evaluation, and how and why they should be involved. You will learn how to monitor evaluation data and engage in process (formative) evaluation as well as product (summative) evaluation.

Chapter Four addresses a number of questions regarding the evaluator’s program description: Why do you need it? How does it relate to your objectives and activities? The chapter will include discussion of the importance of gleaning information from program staff, for example, goals and objectives, the activities planned to achieve the goals and objectives, and the measures that will be used to evaluate the results. The discussion may also involve more than one level of goals and objectives, such as those of the organization, the staff members, and the clients.

The purpose of Chapter Five is to describe different models of evaluation and to illustrate the circumstances under which each is used. Certain design components are shared by all models: evaluation questions, activities to observe, data sources, whom you collect data from or about, how and when you collect the data, how you analyze the data, and who is responsible for each piece of the evaluation.

In Chapter Six, you will learn to identify the data sources that will help you, as an evaluator, to determine whether the program has met its objectives. Two important questions are answered: How do data sources relate to the evaluator's program description? Are paper-and-pencil instruments the only form of measurement? This chapter also introduces logic models as an additional means of understanding project concept and project flow. If a logic model exists, it is advisable that an evaluator discusses it with the project staff so that there is total understanding of the intent, process, and intended outcomes. If one does not exist, it might be helpful for the evaluator to construct one with the project staff so that there is understanding and agreement among all stakeholders of the project staffs' intent, process, and intended outcomes.

Chapter Six discusses many forms of data collection, among them using existing data, such as previously collected public records, and collecting new data using surveys, checklists, interview schedules, objective tests and scales, projective measures, and observational analysis.

Chapter Six also differentiates between data collection formats that are obtrusive (when people know you are collecting information) and those that are unobtrusive (when people are unaware that you are collecting information).

In Chapter Seven, data analysis will become an understandable concept for you. Specifically, Chapter Seven introduces data analysis by levels of measurement: nominal, ordinal, interval, and ratio. Data analysis terms such as measures of central tendency and analysis of variance, which are ways of analyzing differences (growth and change), are discussed by level.

The purpose of Chapter Eight is to help you distinguish between evaluation and research. The chapter discusses both quantitative and qualitative approaches. You will also learn the essential facts about samples, including why you would use sampling, what the various sampling terms mean, and what you need to know about sample size. These questions and others are answered in very clear language: What are control and experimental groups? Is it always possible to have these in your evaluation? When is it possible, and when is it not? The chapter also addresses how you deal with nonresponse.

Chapter Nine, Writing the Evaluation Report, has its basis in Chapter Three's learning about the program's stakeholders. All or any of the stakeholders might receive the report. The chapter reviews your audiences: clients, program staff, organization, or sponsor. In addition, the chapter suggests the many possibilities for focus in the report. As the report writer, you may focus on any or all of the following points: philosophy and goals of the organization, client, community; the needs assessment; program planning; program implementation; and evaluation. Chapter Nine suggests an outline that an evaluator can adopt or adapt for his or her own report.

The Appendix, More on Data Analysis, is included for those readers who wish to delve further into the specifics of data analysis.

Finally, we present a sample evaluation report that has been annotated to illustrate the concepts in this book.

Using This Book

The ABCs of Evaluation is applicable to all sectors of the organizational world (that is, business, education, community-based organizations, the health sector, the public sector) rather than focusing on any one sector. Note that the book will not provide specific instructions on designing personnel appraisals. However, by becoming familiar with the basic steps in the evaluation process, following the scenarios provided in each

chapter, and using the additional resources at the end of each chapter, the reader should become considerably more comfortable with the evaluation process. The novice evaluator will find it an invaluable guide leading to yet more lessons that can only be learned through the experience of conducting both internal and external evaluations.

Depending on your level of experience, knowledge, or comfort with the evaluation processes, you may wish to cover the entire text in a stepwise progression or use parts of it to complete and complement what you already know. Whatever the case, *The ABCs of Evaluation* can take on the role of primer for the novice evaluator or resource for the seasoned evaluator.

THE AUTHORS

John Boulmetis is a professor in the School of Education (SOE), director of Graduate Studies for the SOE, codirector of the URI/RIC PhD Program, and coordinator of the graduate program in Adult Education at the University of Rhode Island. Since 1985 he has also been director of URI's Center for Human Services (previously the Institute of Human Science and Services). Dr. Boulmetis received his master's degree in education from the University of Rhode Island, and his doctoral degree in adult and vocational education from Ohio State University. During the past thirty-eight years, Dr. Boulmetis has been consultant to public and private institutions for the purposes of training, training development, third-party evaluation, and program planning. He has assisted in the development, operation, and conduct of a national study of competency-based adult vocational programs in the United States, and has been principal investigator on projects dealing with evaluation, needs assessment, curriculum development, and training within human services agencies in the states of Rhode Island and Connecticut. He is a member of the American Association of Adult and Continuing Education (AAACE), and the Association of Career and Technical Education, and was president of the AAACE in 2000. Dr. Boulmetis was honored in 2005 and 2009 by the AAACE with the President's Award

for Exceptional and Innovative Leadership in Adult and Continuing Education. In 1987 he was honored by the Rhode Island Association of Adult and Continuing Education as the Rhode Island Adult Educator of the Year. He is the author of numerous keynote addresses at national and international conferences, evaluation reports, articles, and books, including *Job Competency: Adult Vocational Instruction* (Fearon Education, 1981).

Phyllis Dutwin is president of Dutwin Associates, a training and development firm, and former vice president of Reading and Educational Services of New Jersey. She received her master's degree in adult education, training, and development from the University of Rhode Island. Dutwin consults with major publishers to develop and write books that help students and adults advance their personal and workplace skills, and has also consulted with the State Department of Education in New Jersey, Texas Instruments, and Textron Chamber of Commerce School in Providence, Rhode Island. An adult educator with more than thirty years of experience in designing reading, writing, and ESL programs, as well as in teaching adults in corporate, classroom, and individual settings, she was honored in 1993 for her outstanding contributions to the Adult Educators of New England. Dutwin is a member of the American Association for Adult and Continuing Education, and the Rhode Island Literacy Adult Council. Among her books are *E-Z ENGLISH* (Barron's, 2009); *English Grammar Demystified* (McGraw-Hill, 2010); *TABE: Tests of Adult Basic Education, Level A* (2003) and *Level D* (2005), *TABE Level A Verbal Workbook* (2008), all for McGraw-Hill; *Grammar in Plain English, 4th edition* (2005), *English the Easy Way, 3rd edition* (2004), *Writing the Easy Way, 3rd edition* (2000), all for Barron's Educational Series; *Read to Work: Business Occupations* (1999) and *Read to Work: Health Occupations* (1999), for Simon and Schuster.

1

What Is Evaluation?

SCENARIO ONE

The administration at Grandview Retirement and Nursing Facility strongly promotes ongoing education and training for all staff. In fact, when new personnel are hired, they agree to take two courses each year, selected from a long list of possible courses decided upon by administration and staff. Topics are many and diverse, from the law and patient rights to medication in geriatric care to the mind-body connection. Two of the recent training programs were devoted to answering these questions: (1) Is there a connection between mental and physical activity and sustained good health and longevity? (2) How do you encourage residents to stay both mentally and physically active? Program and nursing staff alike found the answer to the first question compelling: activity is indeed linked to well-being and longevity (Wells, 1997). They decided to institute additional programs for residents that would involve mental and physical activity.

Coincidentally, a volunteer had been coming to the retirement and nursing facility for about six months to work with interested residents in an informal gardening program. The volunteer, Ruth, discovered that one of the things the residents missed most when they left their homes was their gardening. The activity helped them regain what they had enjoyed and gave them a real opportunity for mental and physical engagement. The facility had a limited budget, so Ruth made her own in-kind donations.

From the beginning, however, the activity was extremely popular and she knew that available resources would not be sufficient to serve all the people who wanted to take part in the program. Ruth volunteered to look for funds, a grant that would allow Grandview to continue and enlarge the program. The administrators said they would be delighted if she would do the legwork, but of course they would need to know exactly what she was doing along the way so that they could keep their board of directors apprised.

A short while later Ruth discovered that the Beed Foundation funded this kind of program and she got their grant application. One of the first items she saw on the application asked for an evaluation of the program. The items in the evaluation question included listing program objectives (for example, to increase residents' activity and mobility level) and listing measurements (both quantitative and qualitative) that would be used as indicators of achievement of those objectives. Ruth later learned that a quantitative measure might be the number of times residents took part in the gardening activities before the onset of the funded program compared to how many at the end of the program. A qualitative measure might be perceptions of garden activity staff regarding the focus and mood of residents.

The grant also called for explaining data collection methods, a discussion of any sampling that would be done, a description of the

evaluation design, the data analysis that would occur, the staffing, and the final report.

Although she was puzzled about most of the items, Ruth hoped that subsequent meetings and discussions with the staff would answer her questions. She needed some basic questions answered first: Why do programs need to be evaluated? What is evaluation? What are you looking for when you evaluate a program? How is this different from research?

Unlike Ruth, readers of this book will not have to wait to have these questions answered. You just need to read Chapter One. After reading this chapter, you should be able to answer the following questions as they relate to the preceding scenario as well as to the chapter material:

1. What do all evaluations have in common?
2. How would you characterize the differences in the efficiency, effectiveness, and impact of a program?
3. Why evaluate in the first place?

Common Denominators

Before we embark on a definition of evaluation, we need to answer the volunteer's question. How does program evaluation differ from research? Unlike evaluation, research takes place in a precisely controlled environment. The Grandview gardening program could not comply with this definition. Among a number of other important reasons that we will discuss in Chapters Five and Eight, Grandview's project staff could not precisely control—nor would they want to—the number and

choice of participants. In addition, research collects data to stretch the envelope of what is known in order to prove or disprove a hypothesis or presupposition. Evaluation looks at program and project objectives and asks whether they have been achieved, judges the worth of ongoing programs, decides upon the usefulness of new programs or projects, and so forth (Rossi and Freeman, 1993).

In both for-profit and nonprofit organizations, managers possess data (information) that could help to evaluate a program or project. These data are the one thing that all evaluations have in common regardless of the particular definition of evaluation one embraces: evaluation is the systematic process of collecting data that help identify the strengths and weaknesses of a program or project. The data may be as simple as records of attendance at training sessions or as complex as test scores showing the impact of a new educational program on increasing students' knowledge across an entire school system.

Whatever definition you apply to evaluation, when your task is to perform a program evaluation you will almost certainly include a number of common evaluation steps. You will find these listed as program cycles in the Putting It All Together section later in this chapter.

Two Definitions of Evaluation

People do not always agree on one definition of evaluation. Following are two different definitions:

- Evaluation is the systematic process of collecting and analyzing data in order to determine whether and to what degree objectives have been or are being achieved.
- Evaluation is the systematic process of collecting and analyzing data in order to make a decision.

Notice that the first ten words in each definition are the same. However, the reasons—the why—for collecting and analyzing the data reflect

a notable difference in the philosophies behind each definition. The first reflects a philosophy that as an evaluator, you are interested in knowing only if something worked, that is, whether it was effective in doing what it was supposed to do. The second statement reflects the philosophy that evaluation makes claims on the value of something in relation to the overall operation of a program, project, or event. Indeed, many experts agree that an evaluation should not only assess program results but also identify ways to improve the program evaluated (Wholey, Hatry, and Newcomer, 1994). A program may be effective but of limited value to the client or sponsor. One can imagine, however, using an evaluation to make a decision (the second definition) even if a program has reached its objectives (the first definition). Federal grants are based on the first statement, that is, whether the program has achieved its objectives, but the harder decision to downsize or change may be a consequence of the second definition of evaluation.

Evaluating Efficiency, Effectiveness, and Impact

We can define evaluation even more precisely as a process that is guided by the reason for doing the evaluation in the first place. An evaluation might be a process of examining a training program in light of values or standards for the purpose of making certain decisions about the efficiency, effectiveness, or impact of the program. To carry out this task, you need to understand the concepts of efficiency, effectiveness, and impact. These three terms will be referred to from this point on as the levels of program evaluation. (See Table 1.1 later in this chapter.)

Efficiency relates to an analysis of the costs—dollars, people, time, facilities, materials, and so forth—that are expended as part of a program in comparison to either their benefits or effectiveness. How is efficiency, or the competence with which a program is carried out, measured in a program? The term itself gives clues to what this is about.

Program monitors look at the efficiency with which details are carried out in a program. Programs often begin with recruiting, gathering materials, providing for space, setting up fiscal procedures, and so forth. Thus the relationship between the costs and end products becomes the focus of an efficiency evaluation. Although very important, these aspects of efficiency may have no bearing on the program's effectiveness. If the investment in the program or project exceeds the returns, there may be little or no efficiency.

For example, let's consider a nuclear power facility that houses a rather substantial training and staff development enterprise. As part of this enterprise, ten instructors are responsible for ensuring that five hundred employees are cycled through training every six months, for a minimum of twenty hours of training each cycle. The training revisits the employees' basic knowledge of their jobs and introduces new concepts developed since the last training. The staff development enterprise might work very efficiently by making sure that all employees cycle through in a timely fashion, in small enough groups to utilize the best of what we know about how adults learn. The students' time on task is often not enough, however, and many of them do not retain much of what was covered in the training. Thus the program is not effective.

The enterprise may be efficient in that it fully utilizes the time of each of the available trainers, it stays within the parameters of the staff development budget, it keeps employee down time to a minimum, it uses materials and equipment that are available, and it completes the training agenda for the company. Yet there may be an increase in accidents or hazardous incidents because employees are making simple, basic mistakes. The enterprise's training has been efficient but not necessarily effective.

When you look at the effectiveness of your program, you are asking this question: "Did the activities do what they were supposed to do?" Simply put, a program's effectiveness is measured in terms of substantive changes in knowledge, attitudes, or skills on the part of a program's clients. Although the right number of participants may have been recruited

and the best possible site may have been secured, the effectiveness test is this: Did the activities provide the skills to run the new equipment? Did the participants gain the knowledge they need to sell the new mortgage or other banking product?

In another example, the same nuclear power plant's staff development program may conduct a training session on a new procedure to decontaminate after entering a containment area. The trainer may pretest all the employees as they begin their training session. Upon completion, the employees are posttested and the results compared to determine whether their knowledge increased, decreased, or stayed the same. An increase in their knowledge would be an indication that the training was effective—it did what it was supposed to do. Yet two weeks after the training, when one of the employees was back at her job post, a situation arose in which she exited a hazardous area after spending some time checking water flow. She used the older, more comfortable procedure for decontamination and caused a problem that put her and her coworkers at risk. Here is an example of training that was effective—the worker passed all the posttests—but had little impact on changing the behavior of the employee.

Thus the impact that the program has had on the people or organization for which it was planned becomes an important evaluation consideration. Impact evaluation examines whether and to what extent long-term and sustained changes occur in a target population. Has the program or project brought about these desired changes? Are employees using the new procedures? And in other scenarios: Are more people off welfare? Has the program changed a family's life? Do your employees have more job satisfaction?

Evaluators frequently pay too little attention to assessing impact. One reason is that impacts often manifest themselves over time, and program managers have already turned their attention elsewhere before computing this aspect of the evaluation. The actual impact that training in new procedures might have in people's everyday life often takes time

to percolate and evolve. An attempt to collect impact data after allowing for this delay may run into a number of blocks, such as learner turnover (you cannot find them), job or circumstance change (they no longer need to use the skills), or lack of time or resources for the evaluator to conduct these follow-up activities.

Still, program and project sponsors are most interested in impacts. Whether a learner feels satisfied with the training, or the training results in knowledge gain means little to a sponsor or employer if the learning doesn't help the organization.

Evaluating Alternatives

The second philosophical statement that defines evaluation presents it as the process of delineating, obtaining, and providing useful information for the purpose of selecting among alternatives. Thus it may not matter whether the program was efficiently conducted, effective, or had an impact on behavior or functions. Instead, the value of the evaluation is in its being able to compare one activity to another, one program to another, or one employee to another so that decisions can be made in the presence of empirically collected data. Search committees perform this kind of evaluation. In the course of their work, they describe job candidates' strengths, outline previous experiences, and acquire other useful information that makes it possible to choose among a number of candidates. A company planning to adopt and purchase a computer system will perform this kind of evaluation on all the systems it is considering. It will select the one that performs the best given the company's needs and resources.

Identifying Areas to Improve

Finally, there is a third way of defining evaluation as the identification of discrepancies between where a program is currently and where it would

like to be. For example, an organization's marketing department may have as one of its goals at least one face-to-face visit with customers per year. Currently, its sales force sees fewer than half the customers in a year. Records of face-to-face calls indicate the discrepancy between where XYZ Corporation is currently and where the organization wants to be.

Personnel evaluations often take on this definition as well. A new employee's first evaluation may be an example of the first definition, that is, an evaluation against some minimal standard of performance. After this initial evaluation, certain performance goals are set for the employee, either mutually or by the supervisor or team. The next and all the subsequent evaluations of that employee are compared with those performance goals or standards. The discrepancies are identified and remediation strategies are developed.

Each of the individuals identified here who have some interest in or ownership of the evaluation are stakeholders in the organization being evaluated. As such, they have a specific focus on what is evaluated and a particular expectation of how they might use evaluation results. Project directors and project managers might focus their attention on how efficiently the project activities are performed and the effectiveness of these activities in reaching predetermined goals and objectives. Project staff (teachers, trainers, counselors, human resource workers, line staff, and volunteers) might focus their attention on how effective the project activities have been in achieving the short-term objectives of their clients and the impact of the activities in sustaining long-term changes in the client. Parent institution administrators and management might focus their attention on the impact of project activities on any significant changes in how the organization functions and the fundamental products of the organization.

Consequently, each of the stakeholders has specific responsibilities before, during, and after the evaluation. Project managers and directors need to determine what aspects of the project *must* be evaluated as

stipulated by funding sources, accreditation requirements, or strategic plans. They also need to identify aspects of the project that should be evaluated to assist them in determining what (or who) is working, not working, no longer needed, and in need of updating. Also, it is usually their responsibility to appoint or hire the evaluator based upon that person's (or agency's) experience, capability, and reputation.

Project staff need to assist in helping the project managers and evaluator in identifying those aspects of the project that should be evaluated that would result in answers for them to improve their practice and product. Also, staff need to assist in data collection both during and after the project cycle to assure that data are collected in a consistent, rigorous, and relevant manner. They act as the reality filter for evaluators to communicate which data collection practices and instruments are realistic given the work being performed.

Organization administrators need to establish a credible atmosphere for the evaluation. They must communicate to all parties involved that the evaluation activities are meaningful, important, and necessary. This means that from the outset they have to get people to understand that the evaluation is not a punitive exercise, but one that will provide important insights to improving the operation of the organization. This sense of improvement must be maintained throughout the evaluation by informing those involved, making changes if activities have been identified as in need of change, and encouraging staff to use interim evaluation findings to improve their practice. At the end of the program cycle, they must *use* the evaluation findings to alter or maintain project operations, instead of just filing the report or forwarding it to the project sponsor.

Other Levels

There are other levels of evaluation as defined by Kirkpatrick (1994; 2009). These levels refer to the eventual use of the evaluation data and

who might make use of the results. Kirkpatrick's four levels of evaluation are as follows:

Level 1, Reaction: Participant impressions

Level 2, Learning: Learning acquired

Level 3, Behavior: Application of the learning

Level 4, Results: Extent that targeted outcomes occur for the company, agency, or school system

In level 1 you are examining the perceptions of individuals who were directly involved as clients of your program. You are interested in their perceptions of how they benefited from the program, what they thought of the program activities, and how they might use what they gained from the program. Program staff will be particularly interested in this level because the feedback will tell them how their efforts are being perceived and used.

In level 2 you are measuring the effectiveness of the program in doing what it set out to do. At this stage you usually compare a set of standards—criteria, goals, or objectives—with the actual results. You might perform pretesting and posttesting of clients to ascertain their change, or, to see the benefits of program participation, compare one group who received the activities of the program with a group who did not. Here you measure the extent to which individuals acquired the knowledge, skills, and attitudes originally projected. Program staff and sponsors will be particularly interested in the results of this level of evaluation.

In level 3 you are attempting to discover the overall impact of the program on clients: Did the program change their long-term behavior, attitudes, or performance? Were the changes observed in the level 2 evaluation sustainable over time? To what degree did they apply what they learned? Sponsors, program staff, and clients will be interested in this level of evaluation.

In level 4 you want to determine the extent to which the parent institution (sponsor) benefited from the program. Here you examine whether targeted outcomes have resulted due to the activities examined under the first three levels. Here your focus shifts from the relative benefit to the client to the relative benefit to the sponsor. To gauge efficiency, you need to determine whether the program's outcome warranted the sponsor's expenditure of resources. To gauge effectiveness, you are investigating whether the productivity of a certain group of employees increased. And to gauge impact, you need to discover whether the program changed the performance, the product, or the image of the sponsor. By addressing these macro concerns at this level of evaluation, you enable decision makers to see more clearly what programs do to help their parent institution.

Formal Reasons to Evaluate

To know what type of evaluation to use, managers first set out their reasons for undertaking an evaluation. In the public sector, for example, federal or state-sponsored programs demand third-party or internal evaluations; these evaluations are not optional. The evaluation may be required for fiscal purposes, with the program's future funding hinging on the results of the evaluation. Or an evaluation may be needed for comparison purposes to find out which of several methods would be the most effective to continue funding.

Certain questions need to be addressed in a formally mandated evaluation. Questions of efficiency might be: When did you begin the program or project? How much did it cost? How long did it take? Questions of effectiveness might be: What did you do? How well did you do it? What were the outcomes? Questions of impact might be: Did the program influence lives? Did the program add value? Table 1.1 shows sample questions for an evaluation of a project to develop an affirmative action plan.