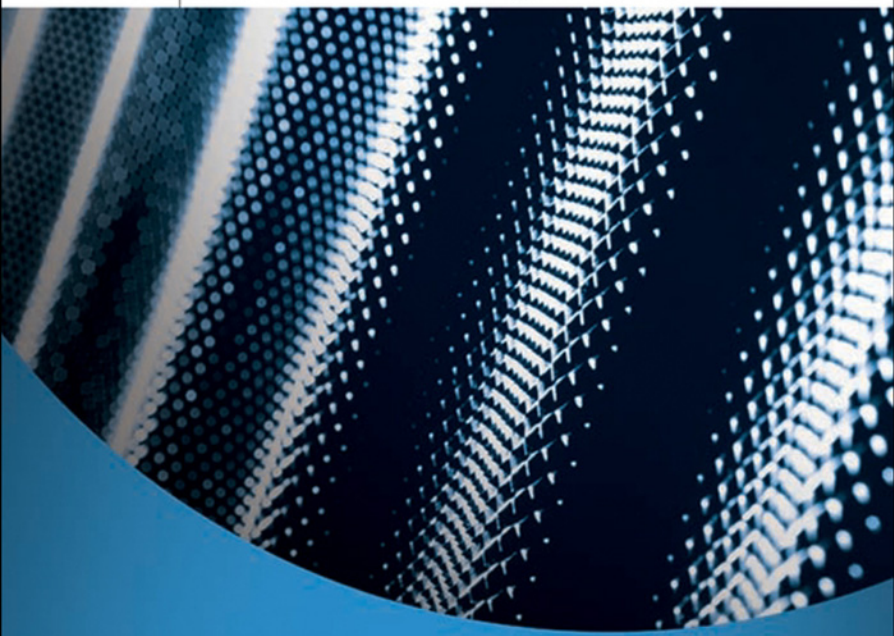
Pfeiffer

Essential resources for training and HR professionals



# DATA COLLECTION

## Planning For and Collecting All Types of Data

PATRICIA PULLIAM PHILLIPS, Ph.D., & CATHY A. STAWARSKI, Ph.D.

MEASUREMENT & EVALUATION SERIES 2

# About This Book

## Why This Book Is Important

This second book in the M&E Series begins with a brief introduction to the ROI process model and the Twelve Guiding Principles. It goes on to explain, in detail, one of the most critical activities in the evaluation process: data collection. Data are collected during every step of the ROI Methodology, and how and when these data are collected are crucial to successful program evaluation. This book provides detailed information about data collection methods and how and when to collect data at each evaluation level.

## What This Book Achieves

This guide details how to develop or use each of the following data collection tools:

- Surveys
- Questionnaires
- Tests
- Simulations
- Interviews
- Focus groups
- Direct observation
- Performance monitoring

- Action plans
- Performance contracts

In addition, this book explains how to select the appropriate data collection method for any situation and any level of evaluation.

## How This Book Is Organized

This book begins by introducing and describing each data collection method and ends by discussing the methods that are best for each evaluation level. Chapter One details how to design effective questionnaires and surveys in order to maximize response rates. It also provides tips for presenting questionnaires to participants in ways that will encourage them to provide valuable data. Chapter Two discusses the major types of tests, including how to develop and administer them. When and how to use simulations to collect data are also explained.

Chapter Three explores how to conduct interviews and focus groups for optimal results, including question development and interview techniques. Guidelines for effective observation are also provided. Chapter Four explains how to use business performance monitoring, action planning, and performance contracts to collect data.

Chapter Five focuses on measuring reaction and planned action and the importance of gathering participant feedback. Chapter Six discusses the reasons for measuring learning data and then describes how to use the data. Measurement and administrative issues involved in data collection at this level are also explored.

Chapter Seven explains the importance of collecting application and implementation data. The challenges and issues of collecting application data are explored, and appropriate data collection methods are also discussed. Chapter Eight explores the critical reasons for collecting impact data. Effective impact measures, appropriate data collection measures, and linking the measures to business needs are all discussed in this chapter.

Finally, Chapter Nine illustrates how to select the appropriate data collection method for each program and each level of evaluation.

# The Measurement and Evaluation Series

Editors

Patricia Pulliam Phillips, Ph.D.

Jack J. Phillips, Ph.D.

# Introduction to the Measurement and Evaluation Series

The ROI Six Pack provides detailed information on developing ROI evaluations, implementing the ROI Methodology, and showing the value of a variety of functions and processes. With detailed examples, tools, templates, shortcuts, and checklists, this series will be a valuable reference for individuals interested in using the ROI Methodology to show the impact of their projects, programs, and processes.

## The Need

Although financial ROI has been measured for over one hundred years to quantify the value of plants, equipment, and companies, the concept has only recently been applied to evaluate the impact of learning and development, human resources, technology, quality, marketing, and other support functions. In the learning and development field alone, the use of ROI has become routine in many organizations. In the past decade, hundreds of organizations have embraced the ROI process to show the impact of many different projects and programs.

Along the way, professionals and practitioners need help. They need tools, templates, and tips, along with explanations, examples, and details, to make this process work. Without this help, using the ROI Methodology to show the value of projects and

programs is difficult. In short, practitioners need shortcuts and proven techniques to minimize the resources required to use this process. Practitioners' needs have created the need for this series. This series will provide the detail necessary to make the ROI Methodology successful within an organization. For easy reference and use, the books are logically arranged to align with the steps of the ROI Methodology.

## Audience

The principal audience for these books is individuals who plan to use the ROI Methodology to show the value of their projects and programs. Such individuals are specialists or managers charged with proving the value of their particular project or program. They need detailed information, know-how, and confidence.

A second audience is those who have used the ROI Methodology for some time but want a quick reference with tips and techniques to make ROI implementation more successful within their organization. This series, which explains the evaluation process in detail, will be a valuable reference set for these individuals, regardless of other ROI publications owned.

A third audience is consultants and re-searchers who want to know how to address specific evaluation issues. Three important challenges face individuals as they measure ROI and conduct ROI evaluations: (1) collecting post-program data, (2) isolating the effects of the program, and (3) converting data to monetary values. A book is devoted to each of these critical issues, allowing researchers and consultants to easily find details on each issue.

A fourth audience is those who are curious about the ROI Methodology and its use. The first book in this series focuses specifically on ROI, its use, and how to determine whether it is appropriate for an organization. When interest is piqued, the remaining books provide more detail.

## Flow of the Books

The six books are presented in a logical sequence, mirroring the ROI process model. Book one, *ROI Fundamentals: Why and When to Measure ROI*, presents the basic ROI Methodology and makes the business case for measuring ROI as it explores the benefits and barriers to implementation. It also examines the type of organization best suited for the ROI Methodology and the best time to implement it. Planning for an ROI evaluation is also explored in this book.

Book two, *Data Collection: Planning For and Collecting All Types of Data*, details data collection by examining the different techniques, processes, and issues involved in this process, with an emphasis on collecting post-program data. It examines the different data collection methods: questionnaires, interviews, focus groups, observation, action plans, performance contracts, and monitoring records.

Book three, *Isolation of Results: Defining the Impact of the Program*, focuses on the most valuable part of the ROI Methodology and the essential step for ensuring credibility. Recognizing that factors other than the program being measured can influence results, this book shows a variety of ways in which the effects of a program can be isolated from other influences. Techniques include comparison analysis using a control group, trend line analysis and forecasting methods, and expert input from a variety of sources.

Book four, *Data Conversion: Calculating the Monetary Benefits*, covers perhaps the second toughest challenge of ROI evaluation: placing monetary value on program benefits. To calculate the ROI, data must be converted to money, and *Data Conversion* shows how this conversion has been accomplished in a variety of organizations. The good news is that standard values are available for many items. When they are not, the book shows different techniques for converting them, ranging from calculating the value from records to seeking experts and searching databases. When data cannot be

converted to money credibly and with minimum resources, they are considered intangible. This book explores the range of intangible benefits and the necessary techniques for collecting, analyzing, and recording them.

Book five, *Costs and ROI: Evaluating at the Ultimate Level*, focuses on costs and ROI. This book shows that all costs must be captured in order to create a fully loaded cost profile. All the costs must be included in order to be conservative and to give the analysis additional credibility. Next, the actual ROI calculation is presented, showing the various assumptions and issues that must be addressed when calculating the ROI. Three different calculations are presented: the benefit-cost ratio, the ROI percentage, and the payback period. The book concludes with several cautions and concerns about the use of ROI and its meaning.

Book six, *Communication and Implementation: Sustaining the Practice*, explores two important issues. The first issue is reporting the results of an evaluation. This is the final part of the ROI Methodology and is necessary to ensure that audiences have the information they need so that improvement processes can be implemented. A range of techniques is available, including face-to-face meetings, brief reports, one-page summaries, routine communications, mass-audience techniques, and electronic media. All are available for reporting evaluation results. The final part of the book focuses on how to sustain the ROI evaluation process: how to use it, keep it going, and make it work in the long term to add value to the organization and, often, to show the value of all the programs and projects within a function or department.

## Terminology: Programs, Projects, Solutions

In this series the terms *program* and *project* are used to describe many processes that can be evaluated using the ROI Methodology. This is an important issue because readers may vary widely in their perspectives. Individuals involved in technology applications may

**Table I.1 Terms and Applications**

| Term | Example |
|---|---|
| Program | Leadership development skills enhancement for senior executives |
| Project | A reengineering scheme for a plastics division |
| System | A fully interconnected network for all branches of a bank |
| Initiative | A faith-based effort to reduce recidivism |
| Policy | A new preschool plan for disadvantaged citizens |
| Procedure | A new scheduling arrangement for truck drivers |
| Event | A golf outing for customers |
| Meeting | A U.S. Coast Guard conference on innovations |
| Process | Quality sampling |
| People | Staff additions in the customer care center |
| Tool | A new means of selecting hotel staff |

use the terms *system* and *technology* rather than *program* or *project*. In public policy, in contrast, the word *program* is prominent. For a professional meetings and events planner, the word *program* may not be pertinent, but in human resources, *program* is often used. Finding one term for all these situations would be difficult. Consequently, the terms *program* and *project* are used interchangeably. Table I.1 lists these and other terms that may be used in other contexts.

## Features

Each book in the series takes a straightforward approach to make it understandable, practical, and useful. Checklists are provided, charts are included, templates are presented, and examples are explored. All are intended to show how the ROI Methodology works. The focus of these books is implementing the process and making it successful within an organization. The methodology is based on the work of hundreds of individuals who have made the ROI Methodology a successful evaluation process within their organizations.

# About Pfeiffer

Pfeiffer serves the professional development and hands-on resource needs of training and human resource practitioners and gives them products to do their jobs better. We deliver proven ideas and solutions from experts in HR development and HR management, and we offer effective and customizable tools to improve workplace performance. From novice to seasoned professional, Pfeiffer is the source you can trust to make yourself and your organization more successful.

**Essential Knowledge** Pfeiffer produces insightful, practical, and comprehensive materials on topics that matter the most to training and HR professionals. Our Essential Knowledge resources translate the expertise of seasoned professionals into practical, how-to guidance on critical workplace issues and problems. These resources are supported by case studies, worksheets, and job aids and are frequently supplemented with CD-ROMs, Web sites, and other means of making the content easier to read, understand, and use.

**Essential Tools** Pfeiffer's Essential Tools resources save time and expense by offering proven, ready-to-use materials—including exercises, activities, games, instruments, and assessments—for use during a training or team-learning event. These resources are frequently offered in looseleaf or CD-ROM format to facilitate copying and customization of the material.

Pfeiffer also recognizes the remarkable power of new technologies in expanding the reach and effectiveness of training. While e-hype has often created whizbang solutions in search of a problem, we are dedicated to bringing convenience and enhancements to proven training solutions. All our e-tools comply with rigorous functionality standards. The most appropriate technology wrapped around essential content yields the perfect solution for today's on-the-go trainers and human resource professionals.

Pfeiffer
www.pfeiffer.com
*Essential resources for training and HR professionals*

# Data Collection

## Planning For and Collecting All Types of Data

Patricia Pulliam Phillips, Ph.D.
Cathy A. Stawarski, Ph.D.

Pfeiffer
A Wiley Imprint
www.pfeiffer.com

Published by Pfeiffer
An Imprint of Wiley
989 Market Street, San Francisco, CA 94103-1741
www.pfeiffer.com



**Library of Congress Cataloging-in-Publication Data**

Phillips, Patricia Pulliam.
Data collection: planning for and collecting all types of data / Patricia Pulliam Phillips, Ph.D., and Cathy A. Stawarski, Ph.D.
p. cm.—(The measurement and evaluation series)
Includes bibliographical references and index.
ISBN: 978-0-7879-8718-3 (pbk.)
1. Project management—Evaluation. 2. Rate of return. 3. Employees—Rating of. 4. Organizational learning—Evaluation. 5. Social sciences—Research—Methodology. I. Stawarski, Cathy A. II. Title.
HD69.P75P496 2008
001.4′33—dc22

2007035758

Production Editor: Michael Kay
Editor: Matthew Davis
Editorial Assistant: Julie Rodriguez
Manufacturing Supervisor: Becky Morgan

Printed in the United States of America

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

# Contents

# Acknowledgments from the Editors


## From Patti

No project, regardless of its size or scope, is completed without the help and support of others. My sincere thanks go to the staff at Pfeiffer. Their support for this project has been relentless. Matt Davis has been the greatest! It is our pleasure and privilege to work with such a professional and creative group of people.

Thanks also go to my husband, Jack. His unwavering support of my work is always evident. His idea for the series was to provide readers with a practical understanding of the various components of a comprehensive measurement and evaluation process. Thank you, Jack, for another fun opportunity!

## From Jack

Many thanks go to the staff who helped make this series a reality. Lori Ditoro did an excellent job of meeting a very tight deadline and delivering a quality manuscript.

Much admiration and thanks go to Patti. She is an astute observer of the ROI Methodology, having observed and learned from hundreds of presentations, consulting assignments, and engagements. In addition, she is an excellent researcher and student of the process, studying how it is developed and how it works. She has become an ROI expert in her own right. Thanks, Patti, for your many contributions. You are a great partner, friend, and spouse.

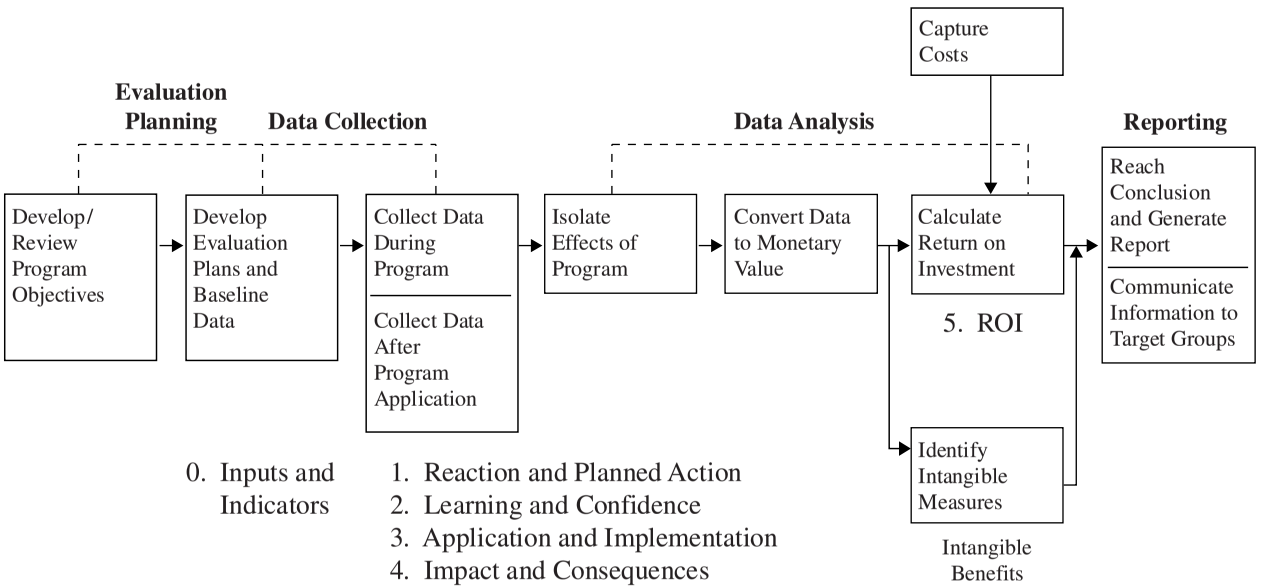
# Principles of the ROI Methodology

The ROI Methodology is a step-by-step tool for evaluating any program, project, or initiative in any organization. Figure P.1 illustrates the ROI process model, which makes a potentially complicated process simple by breaking it into sequential steps. The ROI process model provides a systematic, step-by-step approach to ROI evaluations that helps keep the process manageable, allowing users to address one issue at a time. The model also emphasizes that the ROI Methodology is a logical, systematic process that flows from one step to another and provides a way for evaluators to collect and analyze six types of data.

Applying the model consistently from one program to another is essential for successful evaluation. To aid consistent application of the model, the ROI Methodology is based on twelve Guiding Principles. These principles are necessary for a credible, conservative approach to evaluation through the different levels.

1. When conducting a higher-level evaluation, collect data at lower levels.
2. When planning a higher-level evaluation, the previous level of evaluation is not required to be comprehensive.
3. When collecting and analyzing data, use only the most credible sources.

Figure P.1. The ROI Process Model



4. When analyzing data, select the most conservative alternative for calculations.
5. Use at least one method to isolate the effects of a project.
6. If no improvement data are available for a population or from a specific source, assume that little or no improvement has occurred.
7. Adjust estimates of improvement for potential errors of estimation.
8. Avoid use of extreme data items and unsupported claims when calculating ROI.
9. Use only the first year of annual benefits in ROI analysis of short-term solutions.
10. Fully load all costs of a solution, project, or program when analyzing ROI.
11. Intangible measures are defined as measures that are purposely not converted to monetary values.
12. Communicate the results of the ROI Methodology to all key stakeholders.

# 1

# Using Questionnaires and Surveys

Data collection is the first operational part of the ROI process model. Data are collected in different time frames and from different sources. This is the first of four chapters on data collection methods. Collectively, these four chapters will provide a variety of ways to meet any application, budget, or time constraint.

Probably the most common data collection method is the questionnaire (Alreck and Settle, 1995). Ranging from short reaction forms to detailed follow-up tools, questionnaires can be used both to obtain subjective information about participants and to document objective, measurable impact results for an ROI analysis. Because of this versatility, the questionnaire is the preferred method for capturing data at Levels 1, 2, 3, and 4 in some organizations.

A survey is a specific type of questionnaire with several applications in measuring program success. Surveys are used in situations in which only attitudes, beliefs, and opinions are captured; questionnaires are much more flexible, capturing a wide range of data from attitudes to specific improvement statistics. The principles of survey construction and design are similar to those of questionnaire design. This chapter explains how to develop both types of instruments.

## Types of Questions

In addition to the types of data sought, the types of questions distinguish surveys from questionnaires. Surveys may elicit

yes-or-no responses, if absolute agreement or disagreement is required, or they may solicit a range of responses, often on a five-point scale from "strongly agree" to "strongly disagree."

A questionnaire may contain any or all of these types of questions:

- *Open-ended questions* allow unlimited answers. Questions are followed by ample blank space for the responses.
- *Checklists* provide a list of items, and the participant is asked to check those that apply in the situation.
- *Two-way questions* limit answers to a pair of alternative responses, such as yes and no.
- *Multiple-choice questions* provide several possible answers, and the participant is asked to select the one that is most applicable.
- *Ranking scales* require the participant to rank a list of items.

## Questionnaire Design Steps

Nothing is more confusing, frustrating, and potentially embarrassing than a poorly designed or improperly worded questionnaire. Fortunately, with thought and planning, these problems can be easily avoided. Questionnaire design is a logical process that can be divided into simple steps. Use the following steps to help you develop a valid, reliable, and effective instrument (Robson, 2002).

### Determine the Specific Information Needed

The first step in questionnaire design is reviewing the objectives, topics, skills, or attitudes presented in the program for potential questionnaire items. Developing this information in outline form is sometimes helpful so that related questions or items can be grouped.

At this time, also explore issues related to the application and impact of the program for inclusion in the questionnaire.

### Involve Stakeholders in the Process

To the extent possible, stakeholders—clients, sponsors, supporters, or other interested parties—should be involved in the questionnaire design process. Ask those most familiar with the program to provide information on specific issues and concerns that might affect how the actual questions are framed for the questionnaire. In some cases, stakeholders may want to provide input on specific issues or items. Not only is stakeholder input useful in questionnaire design but it also builds ownership in the measurement and evaluation process and supports content validity.

### Select the Types of Questions

From the five types of questions described previously, select the type or types that will result in the specific data needed. The planned data analysis and variety of data needed should be considered when deciding which types of questions to use.

### Develop the Questions

The next step is to develop specific questions based on the type of questions selected and the information needed. Questions should be simple and straightforward, to avoid confusing the participants or leading them toward a desired response. Each question should address only one issue. If multiple issues need to be addressed, divide questions into multiple parts or develop separate questions for each issue. Avoid terms or expressions that might be unfamiliar to participants.

### Check the Reading Level

To ensure that the questionnaire can be easily understood by the target audience, assess the reading level of the questionnaire. Many word processing programs have a function that can determine the

reading difficulty of a text, indicating what grade level of education would be needed to read it. This important check ensures that the reading level of the questionnaire matches that of the target audience.

### Test the Questions

Proposed questions should be tested to make sure that they will be correctly understood. Ideally, the questions should be tested on a sample group of participants. If this is not feasible, the sample group of employees should be at approximately the same job level as the participants. Seek feedback, critiques, and suggestions from the sample group so that the questionnaire design can be improved before it is administered to participants. Ensure that questions reflect program objectives and content.

### Address the Anonymity Issue

Participants must feel free to respond openly to questions, without fear of reprisal. The confidentiality of their responses is of the utmost importance because there is usually a link between a questionnaire's anonymity and respondents' honesty. Therefore, questionnaires and surveys should be anonymous unless individuals must be identified for specific reasons. In situations in which participants must complete the questionnaire as a captive audience or submit a completed questionnaire directly to an individual, have a neutral third party collect and process the data, to ensure that participants' identities are not revealed. In cases in which individual identities must be known (for example, to compare output data with previous data or to verify the data), make every effort to keep respondents' identities from being revealed to those who might be biased by their responses. Confidentiality goes a long way when collecting data for current or future evaluation projects.

### Design for Ease of Tabulation and Analysis

Consider how each potential question will affect data tabulation, data summary, and data analysis. If possible, outline and review

the data analysis process at this point. This step will help you avoid problems of inadequate, cumbersome, or lengthy data analysis caused by improper wording or design of questionnaire items.

### Develop the Completed Questionnaire and Prepare a Data Summary

Integrate and develop the questions into an attractive questionnaire with instructions that will allow it to be administered effectively. In addition, develop an analysis spreadsheet so that the data can be tabulated quickly for analysis. Developing the questionnaire and planning data analysis in tandem will ensure appropriate and efficient reporting in the end.

## Improving the Response Rate for Questionnaires and Surveys

The items on a questionnaire represent a wide range of potential issues to explore. Obviously, asking all of the possible questions could result in a reduced response rate. The challenge, therefore, is to design and administer the questionnaire so as to maximize the response rate while ensuring reliable responses. Asking too many questions can reduce the number of questionnaires returned; asking too few can negatively affect the reliability of the results. Response rate management is a critical issue if the questionnaire is the primary data collection method and most of the evaluation hinges on the response of the participants. Taking the actions discussed in this section can help increase response rates.

### Provide Advance Communication

If it is appropriate and feasible, communicate with participants in advance about the requirement of completing the questionnaire. Advance warning reduces some of the resistance to the process, provides an opportunity to explain in more detail the circumstances surrounding the evaluation, and positions the follow-up evaluation as an integral part of the program, not an add-on activity.

### Communicate the Purpose

Make sure that participants understand the reason for the questionnaire, including who or what has initiated this specific evaluation and how the data from the questionnaire will be used. Participants should know whether the questionnaire—and the evaluation it is a part of—is the result of an ongoing systematic process or a special request for this program. Let the participants know who will see the data and the results of the questionnaire. If the questionnaire is anonymous, communicate clearly to participants the steps that will be taken to ensure anonymity. If senior executives will see the aggregate results, let participants know.

### Describe the Data Integration Process

If other data are being collected for the same evaluation, help participants understand how the questionnaire results will be combined with those other data. The questionnaire may be only one of several data collection methods used. Participants should know how the data will be weighted and integrated into the final report.

### Keep the Questionnaire as Simple as Possible

While a simple questionnaire does not always provide the full scope of data necessary for an ROI analysis, a simple approach should be a goal. When questions are developed and the total scope of the questionnaire is finalized, keep it as simple and brief as possible. But, as mentioned earlier, take care to design a questionnaire that provides reliable results as well as one that will ensure high response.

### Simplify the Response Process

To the extent possible, responding to the questionnaire should be easy. If appropriate, a pre-addressed, stamped envelope should be included. Perhaps e-mail can be used to respond, if that is easier. In some situations, Web-based questionnaires are better. In still other

situations, a response box can be provided near the participants' workstations.

### Use Local Manager Support

Management involvement at the departmental or even functional level is critical to response rate success. Ask participants' managers to distribute the questionnaire, refer to the questionnaire in staff meetings, follow up to see whether the questionnaire has been completed, or generally show support for completing the questionnaire. Direct manager support will cause some participants to respond with usable data.

### Let the Participants Know That They Are Part of a Sample

If it is appropriate, let participants know that they are part of a carefully selected sample and that their input will be used to make decisions that will affect a much larger target audience. This knowledge often appeals to participants' sense of responsibility, motivating them to provide usable, accurate data on the questionnaire.

### Consider Incentives

A variety of incentives can be offered. Usually, incentives fall into three categories. The first type of incentive is provided in exchange for the completed questionnaire. For example, if participants return the questionnaire personally or through the mail, they receive a small gift, such as a mouse pad or coffee mug. If anonymity is an issue, a neutral third party can distribute the gifts.

The second type of incentive is provided to make participants feel guilty about not responding. Examples are a dollar bill (or equivalent international currency) clipped to the questionnaire or a pen enclosed in the envelope. Participants are asked to "take the money, buy a beverage, and fill out the questionnaire" or to "please use this pen to complete the questionnaire."

The third type of incentive is designed to obtain a quick response. This approach is based on the assumption that quick

responses will ensure a greater response rate. If an individual puts off completing the questionnaire, the odds of completing it diminish. The initial group of participants who send in the questionnaire may receive a more expensive gift, or members of this group may be entered in a drawing for a gift. For example, in one study involving seventy-five participants, the first twenty-five who returned questionnaires were placed in a drawing for a $500 gift certificate. The next twenty-five were added to the first twenty-five for another drawing. After the first fifty, there was no incentive. The longer a participant waited, the lower his or her odds of winning were.

**Have an Executive Sign the Introductory Letter**

Participants are always interested in who signed the letter that accompanies a questionnaire. For maximum effectiveness, a senior executive who is responsible for a major area in which the participants work should sign the letter. Employees may be more willing to respond to a senior executive's request than a staff member's.

**Use Follow-Up Reminders**

Send a follow-up reminder one week after participants receive the questionnaire and another reminder two weeks after the questionnaire is received. In some situations, a third follow-up is recommended. Don Dillman (2006) suggests a tried-and-true approach that includes four follow-ups to an initial invitation to participate in the survey. Sometimes, it is effective to send the follow-ups through different media. For example, a questionnaire might be sent through the regular mail, the first follow-up reminder might be a telephone call, and a second follow-up reminder might be sent through e-mail.

**Provide a Copy of the Results to the Participants**

Make sure that participants see the results of the study, even if in an abbreviated form. More important, tell participants at the time they are asked to provide the data how and when they will receive

a copy of the study. This promise often increases the response rate because some individuals want to see the results for the entire group.

### Review the Questionnaire with Participants

The participants must understand the questionnaire. Seeing a copy in advance of the data collection can be helpful to them. Ideally, the questionnaire should be distributed and reviewed at the program launch or during the initial project meeting. Each question should be briefly discussed, and any issues or concerns about the questions should be clarified. This not only helps the response rate but also improves the quality and quantity of data.

### Consider a Captive Audience

The best way to obtain a high response rate is to collect data from a captive audience. As part of a program follow-up session, a routine meeting, or a mandatory session designed for data collection, ask participants to provide input, usually during the first few minutes of the meeting. Sometimes, a routine meeting (such as a weekly sales meeting or staff meeting) provides the perfect setting for collecting data. This approach is ideal in a major program with a series of meetings; in that case, each subsequent meeting is an opportunity to collect data about the previous one.

### Communicate the Timing of Data Flow

Give participants a specific deadline for completing the questionnaire. In addition, let them know when the results will be available. The best approach is to provide the date when the last questionnaires will be accepted and the date they will receive the results of the evaluation. It may also be prudent to advise participants when the sponsor will receive the results so that they can anticipate potential feedback, changes, or opportunities. Informing participants about specific points on the timetable builds respect for the entire evaluation process. This commitment of a timetable also builds accountability into management of the evaluation project.

### Select the Appropriate Medium

It is important that the medium of the questionnaire (for example, paper, Web, or e-mail) match the culture of the participants. The medium should be selected for the convenience of the respondents, not the evaluator. Sometimes, an optional response medium can be offered in order to make responding more convenient for some participants.

### Consider Anonymous or Confidential Input

Anonymous data are often more objective and, sometimes, more free-flowing than data not provided anonymously. If participants know that their input is anonymous, they will be more constructive and candid in their feedback, and their response rates will, in most situations, be higher. Sometimes, however, it is important to know who is responding. When this is the case, it is still important to ensure confidentiality. A confidentiality statement indicating that participants' names will not be revealed to anyone other than those collecting and analyzing data can play a positive role in gaining a high response rate. An explanation of how the confidentiality will be managed reinforces the commitment to keeps names separate from data.

### Pilot Test the Questionnaire

Consider running a pilot test on a sample of the target audience. Conducting a pilot test is one of the best ways to ensure that a questionnaire is designed properly and that the questions flow adequately. Pilot testing can be accomplished quickly with a very small sample and can reveal problems with a questionnaire before it is administered to the whole audience. This will alleviate potential confusion, which sometimes negatively influences participants' willingness to respond.

### Explain How Long Completing the Questionnaire Will Take

Participants need a realistic guideline on how long it will take them to provide the data. Although this issue seems simple, it should not be overlooked. Nothing is more frustrating than grossly underestimating how much time will be needed to complete a questionnaire. If the questionnaire is online, consider placing an indicator in a visible location on the screen in order to give the respondents an idea of how many more questions to expect. The pilot test should provide the information needed to allocate adequate time for participants to respond.

### Personalize the Process

Participants usually respond well to personal messages and requests. Personalize the letter accompanying the questionnaire, if possible. In addition, if it is possible, use personal phone calls to deliver follow-up reminders. Calls may be made by the program facilitator, a manager or supervisor, an executive, or even the expert in the field being introduced to participants. A personal touch brings sincerity and a sense of responsibility to the process. It also further encourages participants to respond by explaining to them the importance of their data.

### Provide an Update

In some cases, it may be appropriate to provide an update on current response totals and the progress of the evaluation project. If individuals understand how others are doing and how many responses have been returned, they may feel subtle pressure and be reminded to provide data by completing their questionnaire.

---

Used together, the steps in this section help boost response rates on follow-up questionnaires. Using all of these strategies can result

in a 50 to 60 percent response rate, even for lengthy questionnaires that take forty-five minutes to complete.

## Final Thoughts

This chapter has briefly described some of the issues involved in using questionnaires and surveys. More detail on questionnaire content will be given in the chapters that discuss the different levels of evaluation and examples of each type of questionnaire.

### References

Alreck, P. L., and Settle, R. B. *The Survey Research Handbook: Guidelines and Strategies for Conducting a Survey*. (2nd ed.) New York: McGraw-Hill, 1995.

Dillman, D. A. *Mail and Internet Surveys: The Tailored Design Method*. (2nd ed.) New York: Wiley, 2006.

Robson, C. *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*. (2nd ed.) Malden, Mass.: Blackwell, 2002.

# 2

# Using Tests

Testing is important for measuring learning in program evaluations. Pre- and post-program comparisons using tests are common. An improvement in test scores shows the change in skills, knowledge, or attitude attributed to the program. The principles of test development are similar to those for the design and development of questionnaires and surveys. This chapter presents information on types of formal testing instruments and on informal tools used to assess knowledge acquisition.

## Types of Tests

Three types of formal tests are commonly used in program evaluation: norm-referenced tests, criterion-referenced tests, and performance tests. In this section, we will examine each of these types of tests.

### Norm-Referenced Tests

Norm-referenced tests compare participants with each other or with other groups rather than against specific instructional objectives. Norm-referenced tests use data to compare the participants with the norm or average. Although in some evaluations, norm-referenced tests can be of only limited use, they may be useful in programs in

which large numbers of participants are involved and average scores and relative rankings are important. In some situations, participants who score highest on the exams receive special recognition or awards or are made eligible for other special activities.

### Criterion-Referenced Tests

A criterion-referenced test is an objective test in which a predetermined cut-off score indicates acceptable performance. A criterion-referenced test is a measure against carefully written objectives for a specific program. In a criterion-referenced test, the interest lies in whether participants meet the desired minimum standard, not in how participants rank against one another. The primary concern is measuring, reporting, and analyzing participant performance as it relates to the instructional objectives. Table 2.1 shows a printout of the results of a criterion-referenced test.

Criterion-referenced testing is a popular measurement method (Shrock and Coscarelli, 2001). Its use is becoming widespread, with frequent use in e-learning. Criterion-referenced tests have the advantages of being objective, precise, and relatively easy to administer. However, their use requires programs with clearly defined objectives that can be measured by tests.

### Performance Tests

Performance testing allows a participant to exhibit a skill or, occasionally, knowledge or attitudes that have been learned in a program. The skill can be manual, verbal, analytical, or a combination of the three. Performance testing is used frequently in job-related training in which the participants need to demonstrate what they have learned. In supervisory and management training, performance testing comes in the form of skill practices or role-playing. Participants are asked to demonstrate the discussion or problem-solving skills that they have acquired.

**Table 2.1. Reporting Format for Data from a Criterion-Referenced Test**

| | Objective 1 | Objective 2 | | | Objective 3 | | | Total Objectives Passed | Minimum Program Standard | Overall Program Score |
|---|---|---|---|---|---|---|---|---|---|---|
| | Pass/Fail | Raw Score | Standard | Pass/Fail | Raw Score | Standard | Pass/Fail | | | |
| Participant 1 | P | 4 | 10 | F | 87 | 90 | F | 1 | 2 of 3 | Fail |
| Participant 2 | F | 12 | 10 | P | 110 | 90 | P | 2 | 2 of 3 | Pass |
| Participant 3 | P | 10 | 10 | P | 100 | 90 | P | 3 | 2 of 3 | Pass |
| Participant 4 | P | 14 | 10 | P | 88 | 90 | F | 2 | 2 of 3 | Pass |
| Totals (4) | 3 Pass, 1 Fail | | | 3 Pass, 1 Fail | | | 2 Pass, 2 Fail | 8 Pass, 4 Fail | | 3 Pass, 1 Fail |

*Source:* Phillips, 1997, p. 124. Used with permission.

To design and administer an effective performance test, as well as other types of tests, follow these recommendations:

- Make sure that the test accurately represents the content of the program and allows the participants to demonstrate as many of the skills that were taught in the program as possible.
- Thoroughly plan every phase of the test, including collection of necessary materials and tools, preparation of the participants, use of time, and evaluation of results.
- Prepare thorough and consistent instructions. As in other types of tests, the quality of the instructions can affect the outcome of a performance test. All participants should be provided with the same instructions.
- Develop standards for a performance test so that participants know in advance what must be accomplished in order for their performance to be considered satisfactory.
- Include relevant information that will keep participants on track and maintain objectivity in responses.

By following these general guidelines, you can develop performance tests into effective tools for program evaluation. Although they are more costly than written tests, performance tests are essential in situations in which the test conditions must mirror the work environment.

## Simulations

Job simulations are another technique for measuring learning. This method involves creating and implementing a procedure or task that simulates or models the activity that the program teaches. The simulation is designed to represent the actual job situation as

closely as possible. Simulations may be used as an integral part of the program's instruction as well as for evaluation. When a simulation is used for evaluation, participants have an opportunity to try performing the simulated activity as well as to have their performance evaluated. Simulations may be used during a program, at the end of a program, or as part of follow-up evaluation. A variety of simulation techniques that are used to evaluate program results will be discussed in this section.

### Electromechanical Simulation

Electromechanical simulation uses a combination of electronic and mechanical devices to simulate real-life situations; this technique is often used in conjunction with programs to develop operational and diagnostic skills.

### Task Simulation

As its name implies, task simulation requires participants, as part of the evaluation process, to perform a simulated task similar to what they would do on the job.

### Business Games

Business games have grown in popularity in recent years. Business games simulate part or all of a business enterprise; participants change the variables of the business and observe the effects of those changes. The game not only reflects the real-world situation but also is a synopsis of the program of which it is a part.

### In-Basket Simulation

The in-basket simulation is particularly useful for assessing learning in supervisory and management training programs. Portions of a supervisor's job are simulated through a series of items that would normally appear in the participant's in-basket, such as memos, notes,

letters, and reports, which simulate real-life conditions that the supervisor will face. The participant's performance in the in-basket simulation represents an evaluation of the program.

### Case Study

Another popular technique is the case study. A case study gives a detailed description of a problem and usually includes a list of several questions. The participant is asked to analyze the case and determine the best course of action. Case studies allow participants to analyze a real-life situation. Although case studies are often ideal for assessing some programs (for evaluating a project or the leadership style of a corporate executive, for example), it is difficult to actually place a score on a person's level of knowledge acquisition. However, case studies do require that participants examine various aspects of program content and transfer that information to the case situation.

### Role-Playing

In role-playing (sometimes referred to as *skill practice*), participants practice a newly learned skill while they are being observed by other individuals. Participants are given specific instructions about their assigned roles, which may include an ultimate course of action. Participants then practice the skill with the other assigned individuals, trying to accomplish the desired objectives and demonstrate the knowledge that they acquired during the program.

---

Simulations come in many varieties. They offer opportunities for participants to practice what was taught in a program and to have those performances observed in conditions similar to those they will face on the job. Job simulations can provide extremely accurate evaluations if objective, clearly measurable criteria are set for performance in the simulation.

## Informal Tests

In some situations, an informal check of learning is necessary to provide some assurance that participants acquired the desired skills, knowledge, or perhaps changes in attitudes as a result of a program. This approach is appropriate when levels of evaluation other than Level 2, learning, are being pursued. For example, if you are planning a Level 3 evaluation of on-the-job application of skills learned in the program, a comprehensive Level 2 evaluation may not be necessary. An informal assessment of learning may be sufficient. After all, resources are scarce and a comprehensive evaluation at all levels can be expensive. The following are some alternative approaches to in-depth learning measurement that might suffice when inexpensive, low-key, informal assessments are needed.

### Exercises, Problems, or Activities

Many programs contain specific activities, exercises, or problems that must be explored, developed, or solved during the program. Some of these are constructed as exercises for group involvement, while others require individual problem-solving skills. When these types of activities are integrated into the program, several specific ways to measure learning are available:

- Submit the results of the exercise for review and evaluation by the facilitator.
- Discuss the results in a group, comparing the different approaches and solutions. Ask the group to assess how much each individual has learned.
- Share the solutions to the problem or exercises, and ask the group or individual participants to provide a self-assessment indicating the degree to which the skills or knowledge have been obtained from the exercise.

- Have the facilitator review the progress or success of individual participants in order to determine their relative success.

**Self-Assessment**

In many applications, a self-assessment may be appropriate. Participants are provided an opportunity to assess the extent of their skills or knowledge acquisition. Self-assessment is suitable when Level 3, 4, or 5 evaluations are planned and knowing whether learning has improved is important. Use the following techniques to ensure that the self-assessment process is effective.

- Allow participants to perform the self-assessment anonymously so that they feel free to provide a realistic and accurate assessment of what they have learned.
- Explain the purpose of the self-assessment, along with the plans for the data. Specifically, discuss the implications for course design or individual testing.
- Explain what the implications of the results will be and how self-assessment and post-program follow-up evaluation may be compared. This may help in reducing the subjectivity and perceived bias that sometimes results when self-assessment evaluations are conducted.

**Facilitator Assessment**

A final informal technique is for facilitators to provide an assessment of the learning that has taken place. Although this is a subjective approach, it may be adequate when a Level 3, 4, or 5 evaluation is planned. One effective way to accomplish such an assessment is to provide a checklist of the specific skills that need to be acquired during the course. Facilitators can then check off their assessment of a participant's knowledge of a particular skill or issue. It may

be appropriate to provide rating scales to guide facilitators in their assessment; a simple yes-or-no checklist is also sufficient if the knowledge area or skill is one that can easily be observed.

## Final Thoughts

This chapter explored a variety of techniques to measure learning. While formal testing provides the greatest opportunity for objective learning assessment, other forms of assessment are also available when tests are inappropriate, unappreciated, or too expensive. Also, it is important to remember that tests provide information on a respondent's success with the test questions. Other forms of assessment used in combination with formal tests will provide a broader and more balanced perspective on success with knowledge and understanding.

### References

Phillips, J. J. *Handbook of Training Evaluation and Measurement Methods*. (3rd ed.) Boston: Butterworth-Heinemann, 1997.

Shrock, S. A., and Coscarelli, W. C. *Criterion-Referenced Test Development*. (2nd ed.) San Francisco: Pfeiffer, 2001.

Westgaard, O. *Tests That Work: Designing and Delivering Fair and Practical Measurement Tools in the Workplace*. San Francisco: Pfeiffer, 1999.

# 3

# Using Interviews, Focus Groups, and Observation

Several methods are available for capturing qualitative data. Three of these—interviews, focus groups, and observation—are the focus of this chapter. Each is a rich and powerful method of data collection.

## Interviews

Although not used in evaluation as frequently as questionnaires, interviews are a useful data collection method. Program staff, the participant's supervisor, or an outside third party can conduct interviews. Interviews can secure data that are not available in performance records or data that are difficult to obtain through written responses or observations (Kvale, 1996). Also, interviews may uncover success stories that can be used when communicating the evaluation results. Participants may be reluctant to list their results on a questionnaire, but they will volunteer the information to a skillful interviewer who asks the right questions and probes for more information. While the interview process uncovers reaction, learning, and impact data, it is primarily used for collecting application data. Interviews may be time-consuming, and they require interviewer preparation to ensure that the process is consistent—both major disadvantages of the process.

### Types of Interviews

Interviews fall into two basic categories: structured and unstructured. A structured interview is much like a questionnaire. Interviewers ask specific questions with little room for deviation from the desired responses. One of the major advantages of structured interviews over questionnaires is that interviews ensure that all the participants respond to all the questions and that the interviewer understands the participant's responses.

An unstructured interview allows the interviewer to probe for additional information. This type of interview uses a few general questions that can lead to more detailed information as important data are uncovered. The interviewer must be skilled in asking follow-up questions and probing for more information when needed.

### Interview Guidelines

The design issues and steps for interviews are similar to those for questionnaires. A few key steps can ensure a successful interview process.

#### *Develop the Questions to Be Asked*

After determining the type of interview that will be used, develop specific questions. The questions should be brief, precise, and designed for easy response. As with questionnaires, interview questions should reflect the objectives of the program.

#### *Test the Interview*

A small number of participants should be interviewed in order to test the interview's design. If possible, these interviews should be conducted as part of the trial run of the program. The responses should be analyzed, and if necessary, the interview should be revised.

*Prepare the Interviewers*

The interviewer should have the appropriate level of core skills, including active listening, asking probing questions, and collecting and summarizing information. Interviewers should also be familiar enough with program content and the evaluation purpose that they can probe for detail when a response infers there is more to the story.

*Provide Clear Instructions to the Participants*

Make sure that the participants understand why the interview is being conducted and that they know how the information will be used. The expectations, conditions, and rules of the interview should also be thoroughly discussed. For example, the participants should know whether their statements will be kept confidential and, if not, who will see them.

*Schedule the Interviews*

Like other evaluation instruments, interviews must be conducted according to a predetermined plan. The timing of the interview, the individual who will conduct it, and its location should all be determined when developing the plan. Commitment from the interviewee should be obtained early. For a program with a large number of participants, interviewing a sample instead of all the participants may be necessary in order to save time and reduce evaluation costs. As with questionnaires, a plan is necessary to ensure that interviewees will follow through on their commitment to participate.

## Focus Groups

Focus groups are a specific type of interview. They are particularly helpful when in-depth feedback is needed for a Level 3 evaluation. A focus group is a small-group discussion conducted by an experienced facilitator and designed to solicit qualitative data on a topic or issue. The basic premise of focus groups is that when

quality judgments are subjective, several individual judgments are better than one. Each group member is required to provide input, for individual input builds on input from others in the group (Subramony, Lindsay, Middlebrook, and Fosse, 2002). Compared with questionnaires, surveys, tests, and interviews, focus groups have several advantages. The group process, in which participants stimulate ideas in others, is an effective method for generating qualitative data. Conducting a focus group is less expensive than carrying out one-on-one interviews, given the longer time it takes to interview the same number of people individually. Focus groups can usually be scheduled quickly, but an incentive needs to be put forth to ensure that targeted respondents participate. The flexible format of a focus group makes it possible to explore a program's unexpected outcomes or applications. Flexibility, however, does not mean lack of structure. The key to a focus group is to keep it focused on a single issue. Otherwise, the focus group will go down many paths, ultimately leaving you with much data, little of which you sought at the outset.

**Applications of Focus Groups**

Focus groups are particularly helpful when qualitative information is needed about a program's success. For example, focus groups can be used in the following situations:

- To evaluate reactions to specific exercises, cases, simulations, or other components of a program
- To assess the overall effectiveness of program application
- To assess the impact of a program in a post-program evaluation

Essentially, focus groups are helpful when evaluation information is needed but cannot be collected adequately with

questionnaires, interviews, or quantitative methods. Situations in which it is helpful for respondents to build off the thoughts of others are good opportunities for focus groups.

### Guidelines

No specific rules have been set on how to use focus groups during an evaluation, though much has been written about various techniques to facilitate and manage focus groups and the resulting data. The guidelines provided in this section can help you plan and conduct a successful focus group.

#### *Plan Topics, Questions, and Strategy Carefully*

As in the implementation of any evaluation instrument, when developing focus groups, planning is critical. The topics, specific questions, and issues to be discussed must be carefully planned and sequenced in order to facilitate the comparison of results of different groups and to ensure that the group process is effective and stays on track.

#### *Keep the Group Size Small*

While there is no magic group size, eight to twelve is appropriate for most focus groups. A group needs to be large enough to ensure different points of view but small enough to provide a chance for each participant to freely exchange comments.

#### *Use a Representative Sample*

If possible, sample groups should be selected to represent the target population. The sample group should match the target population in its mix of job experience, rank, and job level. The best way to select a representative sample is through the random selection process, in which everyone in the population has an equal opportunity to participate; however, this technique is not always feasible.

*Use Experienced Facilitators*

A focus group's success rests with the facilitator, who must be skilled in conducting focus groups. Facilitators must know how to control aggressive members of the group and how to diffuse the input from those who may want to dominate the group. Also, facilitators must be able to create an environment in which participants feel comfortable enough to freely and openly offer comments. To facilitate a free flow of comments, some organizations use external facilitators.

---

In summary, a focus group is an inexpensive and quick way to determine the strengths and weaknesses of a program. However, the data represent only the views of those participating in the focus group. This can limit the perspective being given about a program. Used in combination with a questionnaire, however, a focus group can provide additional information on a specific area covered in the questionnaire, thereby broadening the understanding of program impact.

## Observations

Another potentially useful data collection method is observing participants and recording any changes in their behavior. The observer may be a member of the staff, the participants' supervisor, a member of a peer group, or an external party. The most common and probably the most practical type of observer is a staff member; however, external, third-party observers often offer a more objective view, although they also increase the cost of data collection.

### Guidelines for Effective Observation

Observation is often misused or misapplied to evaluation situations, causing some to lose faith in the process. The effectiveness of

observation can be improved by following the guidelines discussed in this section.

*Observations Should Be Systematic*

The observation process must be planned so that it can be executed effectively and without surprises. The individuals observed should know in advance that they will be observed and why they will be observed, but not necessarily when, unless the observation is used to measure learning or knowledge acquisition. If the observation is used to collect data regarding a participant's routine application of knowledge, the observer should go unnoticed; otherwise, the behavior may be influenced. The timing of observations should be part of the plan as well. There are right times and wrong times to observe a participant. If a participant is observed when times are not normal (such as during a crisis), the data that are collected may be useless.

*Observers Should Be Knowledgeable*

Observations involve judgment decisions. The observers should know how to interpret and report what they see. They must analyze which behaviors are being displayed and what actions the participants are taking. They should know how to summarize the behaviors and report the results in a way that is meaningful to management and senior executives.

*The Observer's Influence Should Be Minimized*

Except for mystery observers and electronic observations, completely isolating the effect of an observer is impossible. If participants know that they are being observed, they may display the behavior they think is appropriate and they will usually be at their best. The presence of the observer must be minimized. To the extent possible, the observer should blend into the work environment or extend the observation period to allow participants to become accustomed to his or her presence.

*Observers Should Be Selected Carefully*

Observers should be independent of the participants—for example, program staff members. An observer from outside the participants' department is often more skilled at recording behavior and making interpretations of behavior than a department person would be and therefore may not need any special preparation. Observers from outside the participating department are usually unbiased in their interpretations. However, an observer from another area of the organization may be perceived as an outsider who is checking up on the work of others. Participants may overreact and possibly resent this kind of observer. Thus, recruiting an observer from outside the organization may sometimes be a better choice. Hiring a consultant has the advantage of avoiding the prejudices that might otherwise affect the observer's decisions; however, as mentioned earlier, this increases the cost of the evaluation. Therefore, when deciding who will observe, a balance must be found between objectivity and cost.

*Observers Must Be Fully Prepared*

Observers must completely understand the information that is needed and what skills were covered during the program. They must be trained for the assignment and provided with a chance to practice their observation skills.

**Observation Methods**

Five methods of observation are used; the choice of method depends on the type of information that is needed. Each method is briefly described in this section.

*Behavior Checklist*

A behavior checklist is useful for recording the presence, absence, frequency, or duration of a participant's behavior as it occurs. A checklist will not usually provide information on the quality or intensity of the observed behavior or on the circumstances surrounding the behavior. Measuring the duration of a behavior is

more difficult and requires a stopwatch and a place on the checklist to record the time interval. Duration, however, is usually not as important as whether a particular behavior was observed or how often it was observed. The number of behaviors listed on the checklist should be small, and they should be listed in a logical sequence if they normally occur in a sequence. A variation of this approach involves coding behaviors on a form, entering a predefined code for each specific behavior. Whether using a checklist or a coded form, the behaviors should be described clearly enough that if multiple observers observe the same person at the same time, they will all give the same response. Ensuring the reliability of the observation is an important issue when using observation methods of any type, particularly checklists.

*Delayed Report*

In the delayed report method, the observer uses no forms or written materials during the observation period. The information is either recorded after the observation is completed or at intervals during the observation. The observer attempts to reconstruct what has been observed during the observation period. The advantage of this approach is that the observer is not as noticeable, and no forms are completed or notes taken during the observation. The observer can blend into the situation and be less distracting. An obvious disadvantage is that information written after the fact may not be as accurate and reliable as information collected at the time that it occurred. A variation of this approach is 360° feedback, a process in which surveys by peers, supervisors, and direct reports are completed based on observations within a specific time frame.

*Video Recording*

Video recording as an observation method has the obvious advantage: the video camera registers every detail of behavior and provides a permanent record. However, this intrusion may be awkward and cumbersome, and participants may be nervous or self-conscious about being videotaped. If the camera is concealed, participants'

privacy may be invaded. For this reason, video recording of on-the-job behavior is not frequently used.

*Audio Monitoring*

Monitoring the conversations of participants who are using the skills they were taught in a program is an effective observation technique. For example, in a large communications company's telemarketing department, sales representatives are trained to sell equipment by telephone. To determine whether employees are using the skills properly, telephone conversations are monitored on a selective and sometimes random basis. While this approach may cause some controversy, it is an effective way to determine whether skills are being applied consistently and effectively. For audio monitoring to work smoothly, it must be fully explained and the rules must be clearly communicated.

*Computer Monitoring*

For employees who work regularly with a keyboard, computer monitoring is becoming an effective way to observe how participants perform job tasks. The computer monitors times, sequences of steps, and other activities to determine whether the participant is performing the work according to what was learned in a program. Because technology continues to be a significant part of many jobs, computer monitoring holds promise for effective observation of the actual application of skills on the job—in other words, Level 3 data.

## Final Thoughts

This chapter discusses three methods of collecting data: interviews, focus groups, and observation. Interviews are similar to questionnaires, except that the interviewer ensures that all the questions are answered and that the answers are understood. In unstructured interviews, the interviewer can ask further questions if needed. Focus groups are helpful for acquiring input from participants when it is important for them to hear what others have to say. Observation

can also be used for data collection, but depending on the observer, quality, objectivity, and costs are considerations.

## References

Kvale, S. *InterViews: An Introduction to Qualitative Research Interviewing*. Thousand Oaks, Calif.: Sage, 1996.

Subramony, D. P., Lindsay, N. R., Middlebrook, R. H., and Fosse, C. E. "Using Focus Group Interviews." *Performance Improvement*, 2002, *41*(8), 38–45.

# 4

# Using Other Data Collection Methods

To round out our description of data collection methods, we now discuss three important approaches: monitoring the records of an organization, which is particularly helpful in collecting impact data; using action plans, which fits many different projects or programs; and performance contracting, which can be very powerful when it is appropriate.

## Business Performance Monitoring

Data that measure performance are available in every organization. Monitoring performance data enables management to measure performance in terms of output, quality, costs, and time. When you are searching for data to use in an evaluation, the first place to look is existing databases and reports. In most organizations, performance data that can be used to show the improvement resulting from a program are available (Mondschein, 1999). If they are not, additional record-keeping systems will have to be developed for measurement and analysis. At this stage, as with many other stages in the process, the question of economics enters. Is it economical to develop the record-keeping system necessary to evaluate a program? Obviously, if the costs are greater than the expected return for the entire program, then developing a new system is not worth the expense.

### Using Current Measures

The recommended approach to collecting performance data is to use existing performance measures, if they are available. Use the guidelines in this section to ensure that analyzing current measures is easy and cost-effective.

*Identify Appropriate Measures*

Performance measures should be researched in order to identify those that are related to the proposed objectives of the program. Frequently, an organization will have several performance measures related to the same item. For example, the efficiency of a production unit can be measured in a variety of ways:

- Number of units produced per hour
- Number of on-schedule production units
- Percentage of equipment used
- Percentage of equipment downtime
- Labor cost per unit of production
- Overtime required per piece of production
- Total unit cost

Each of these, in its own way, measures the efficiency or effectiveness of the production unit. All related measures should be reviewed in order to identify those most relevant to the program.

*Convert Current Measures to Usable Ones*

Occasionally, existing performance measures are integrated with other data. Keeping them isolated from unrelated data can be difficult. In this situation, all existing measures should be extracted and recalculated as necessary to render them more appropriate for comparison in an evaluation. At times, conversion factors may be necessary. For example, the average number of new sales orders per

month may be presented regularly in the performance measures for the sales department. In addition, the sales costs per sales representative may be presented. However, in the course of evaluating a new program, the average cost per new sale is needed. The two existing performance records can be used to develop the data needed for the comparison.

### Developing New Measures

In some cases, data needed to measure the effectiveness of a program are not available. The staff must work with the organization to develop record-keeping systems, if this is economically feasible. In one organization, a new employee orientation system was implemented company-wide. Several measures were planned, including early turnover—represented by the percentage of employees who left the company during the first six months of their employment. An improved employee orientation program that would influence this measure was planned. At the time of the program's inception, this measure was not monitored. When the program was implemented, the organization began collecting early turnover figures for comparison.

Following are some typical questions to consider when creating new measures:

- Which department will develop the measurement system?
- Who will record and monitor the data?
- Where will the measures be recorded?
- Will forms be used?

Answering these questions will usually involve other departments or management decisions beyond the scope of the staff members who are planning the program. An organization's administration division, finance and accounting departments, or information

technology function may be instrumental in helping to determine whether new measures are needed and, if so, how they will be collected.

## Action Planning

The action plan is a versatile follow-up process. In this approach, participants—the individuals directly involved in the program who must apply the objectives of the program on the job—are required to develop action plans as part of the program. Action plans contain detailed steps for accomplishing specific objectives related to the program. The plan is typically prepared on a printed form such as the one shown in Exhibit 4.1. The action plan shows what is to be done, by whom, and by what date the objectives should be accomplished. The action plan approach is a straightforward, easy-to-use method for determining how participants will change their behavior on the job and achieve success as a result of the program. Action plans produce data that answer such questions as

- What steps or action items have been accomplished, and when?
- What on-the-job improvements or accomplishments have been realized since the program was conducted?
- What is the monetary value of the improvement?
- How much of these improvements is linked to the program versus other factors?
- What may have prevented participants from accomplishing specific action items?

With the information provided in the action plan, professionals can decide whether a program should be modified and in what ways, while managers can assess the findings to determine the value of the program.

## Exhibit 4.1. Action Plan

**Name:** _____ **Instructor Signature:** _____ **Follow-Up Date:** _____

**Objective:** _____ **Evaluation Period:** _____ **to** _____

**Improvement Measure:** _____ **Current Performance:** _____ **Target Performance:** _____

| Action Steps | Analysis |
|---|---|
| 1. _____<br>_____<br>2. _____<br>_____<br>3. _____<br>_____<br>4. _____<br>_____<br>5. _____<br>_____<br>6. _____<br>_____<br>7. _____<br>_____ | A. What is the unit of measure?_____<br>B. What is the value (cost) of one unit? $ _____<br>C. How did you arrive at this value?<br>_____<br>_____<br>_____<br>_____<br>D. How much did the measure change during the evaluation period?<br>_____ (monthly value)<br>E. List the other factors that may have caused this change:<br>_____<br>_____<br>F. What percentage of this change was actually caused by this program? _____%<br>G. What level of confidence do you place on the above information? (100% = certainty; 0% = no confidence) _____% |
| Intangible Benefits: _____ | |

**Comments (i.e., barriers or enablers affecting completion of action items):** _____

_____

_____

### Developing an Action Plan

Development of an action plan requires that two tasks be completed during the program: (1) determining the areas for action and (2) writing the action items. The areas or measures for action should stem from the original need for the program and from the content of the program; at the same time, they should be related to on-the-job activities. Participants can independently develop a list of potential areas for action, or a list can be generated during group discussions. The list may include a specific business measure needing improvement, or it may represent an opportunity for increased performance in a certain area. Typical categories are

- Productivity
- Sales, revenue
- Quality or process improvement
- Efficiency
- Time savings
- Cost savings
- Complaints
- Job satisfaction
- Work habits
- Customer satisfaction
- Customer service

The specific action items should support business measures; therefore, writing them is usually more difficult than simply identifying the action areas. The most important characteristic of an action item is that it be written so that everyone involved will understand it and will know when it has been achieved. One way

to help attain this goal is to use specific action verbs. Here are some examples of action items:

- *Learn* how to use Microsoft Vista software by [date].
- *Identify* and *secure* five new customer accounts by [date].
- *Handle* every piece of paper only once, to improve my personal time management, by [date].
- *Learn* to talk with my employers directly about problems that arise rather than avoiding a confrontation, by [date].

Some questions that typically are asked when developing action items are

- How much time will this action take?
- Are the skills to accomplish this action item available?
- Who has the authority to implement the action plan?
- Will this action have an effect on other individuals?
- Are there any organizational constraints on accomplishing this action item?

If appropriate, each action item should have a completion date and should indicate any resources or actions by other individuals that are required for completion. Also, planned behavior changes should be observable. Completion of the action item should be obvious to the participant and to others when it occurs. Action plans, as used in this context, do not require prior approval or input from the participant's supervisor, although that may be helpful in gaining the supervisor's support for the process. Ultimately, it is the participant involved in the program or project who is accountable

for completion of the action plan, though his or her success is sometimes influenced by others.

### Using Action Plans Successfully

The action plan process should be an integral part of a program, not an add-on or optional activity. To obtain maximum effectiveness from action plans and to collect data for ROI calculations, the following steps should be implemented.

*Communicate the Action Plan Requirement Early*

Participants often react negatively to action plans when they are surprised by them. When program participants are asked to develop a detailed action plan that they were not expecting, there is often immediate, built-in resistance. Communicating the need for an action plan to participants in advance and presenting action plans as an integral part of the program will often minimize resistance. When participants fully realize the benefits of formulating an action plan before they attend the first session, they often take the process more seriously and will usually perform the extra steps to make it a success. In this scenario, the action plan is positioned as an application tool, not an evaluation tool.

*Describe the Action Planning Process at the Beginning of the Program*

During the first session, discuss action plan requirements, including an explanation of the purpose of the process, why it is necessary, and the basic participant requirements during and after the program. Some facilitators provide a separate notepad so that participants can collect ideas and useful techniques to use in developing their action plans. This productively focuses more attention and effort on the process.

*Teach the Action Planning Process*

An important prerequisite for successful use of action plans for collecting data is ensuring that the participants understand how an

action plan works and how to develop one. A portion of the program's agenda should be allocated to teaching participants how to develop plans. During this session, the requirements are outlined, special forms and procedures are discussed, and a completed example is distributed and reviewed. Sometimes, an entire program module is allocated to the action planning process so that participants will fully understand it and use it. Any available support tools—such as key measures, charts, graphs, suggested topics, or sample calculations—should be used in this session to help facilitate the plan's development.

*Allow Time to Develop the Plan*

When action plans are used to collect data for an ROI evaluation, participants must be allowed time to develop their plans during the program. Sometimes, it is helpful to have participants work in teams so that they can share ideas as they develop their individual plans. During these sessions, facilitators often monitor the progress of individuals or teams in order to keep the process on track and to answer questions. In some programs, action plans are developed in an evening session as a scheduled part of the program.

*Have the Facilitator Approve Action Plans*

The action plan must be related to program objectives and, at the same time, represent an important accomplishment for the organization when it is completed. Participants may stray from the intent and purpose of action planning and not give it the attention that it requires. Therefore, it is helpful to have the facilitator or program director sign off on the action plan, ensuring that the plan reflects all the requirements and is appropriate for the program. In some cases, a space on the action plan is provided for the facilitator's signature.

*Require Participants to Assign a Monetary Value to Each Improvement*

Participants are asked to determine, calculate, or estimate the monetary value of each improvement listed in the plan. When the actual

improvement occurs, participants can use these values to capture the annual monetary benefits of the plan. For this step to be effective, it may be helpful to provide examples of typical ways in which values can be assigned to the actual data.

*Ask Participants to Isolate the Effects of the Program*

Although the action plan was developed because of the program, the actual improvements resulting from participants' completion of the actions may have occurred as a result of other factors. It is, therefore, imperative that participants recognize the other influencing factors and give credit only for the improvement actually due to their completion of the actions on the action plan. For example, a participant may target a reduction in employee turnover as the measure he or she intends to improve. Specific actions are listed on the action plan based on information or knowledge the participant acquired during the program. When follow-up on the action plan occurs, employee turnover may well have been reduced, but this reduction could have been the result of a number of factors as well as the completion of the action. Factors such as leadership changes or bonus-driven pay plans or new hiring practices can all affect employee turnover. Therefore, when completing the action plan during the follow-up evaluation, participants should isolate the impact of their actions from the impacts of those other influences (Phillips and Phillips, 2002). While at least nine ways to isolate the effects of a program are available, participant estimation is usually appropriate when action plans are used. Consequently, participants are asked to estimate the percentage of the improvement that is directly related to the program. This question can be asked on the action plan or on a follow-up questionnaire.

*Ask Participants to Provide a Confidence Level for Estimates*

Estimates of program impact and monetary values of business impact measures may not be exact. Therefore, participants are asked to indicate their level of confidence in those two values, collectively.

Assigning a value on a scale of 0 to 100 percent (in which 0 percent means no confidence and 100 percent means complete confidence) provides participants with a way to express any uneasiness or lack of confidence in their estimate. This error adjustment results in the most conservative estimate given the range of possibilities based on the participant's input.

*Require That Action Plans Be Presented to the Group*

There is no better way to secure commitment to and ownership of the action planning process than to have a participant describe his or her action plan to fellow participants. Having participants present their action plans helps ensure that the process is thoroughly developed and that it will be implemented on the job. Sometimes, the process causes competition among the group. If the number of participants is too large for individual presentations, perhaps one participant can be selected from each team (if the plans are developed in teams). Under these circumstances, a team will usually select the best action plan for presentation to the group, raising the bar for others.

*Explain the Follow-Up Process*

Participants must leave the session with a clear understanding of the timing of action plan implementation and planned follow-up. How the data will be collected, analyzed, and reported should also be openly discussed. Five options are available for following up on the implementation of action plans:

1. The group is reconvened to discuss the progress on the plans.
2. Participants meet with their immediate supervisor to discuss the success of the plan. Copies of the minutes from those meetings are forwarded to the program team.
3. The program evaluator, the participant, and the participant's manager meet to discuss the plan and how it was implemented.

4. Participants send their plan to the evaluator, and they discuss it in a telephone call.
5. Participants send the plan directly to the evaluation team, with no meetings or discussions. This is the most commonly used option.

Other data collection options are available. Select a method that fits the culture, requirements, and constraints of the organization.

*Collect Action Plans at the Stated Follow-Up Time*

Because having an excellent response rate is critical to the success of data collection through action plans as well as other data collection techniques, it is important to take the necessary steps to ensure that the action plans are completed and the data are returned to the appropriate individual or group for analysis. Once a collection time has been determined, every effort must be made to get the action plans returned to the evaluation team. Some organizations use follow-up reminders by mail or e-mail to encourage participants to complete their action plans. Others periodically call participants to check on their progress. Still others offer assistance in completing the final plan. These steps may require additional resources, however, and these costs must be weighed against the importance of having more data.

When the action plan process is implemented as outlined in this book, response rates will normally be very high: in the 60 to 90 percent range. Usually, participants see the importance of the process, so they develop their plans in detail before leaving the program and complete them at the assigned time.

*Summarize the Data and Calculate the ROI*

If developed and implemented properly, each action plan will have annualized monetary values associated with improvements in the

targeted business measures. Also, each individual will have indicated the percentage of the improvement that is directly related to the program. Finally, each participant will have provided a confidence percentage that reflects any uncertainty about the process and the subjective nature of some of the data provided.

Because this process involves estimates, it may not appear credible. Several adjustments can be made during the analysis to improve the credibility of the process. Make adjustments by following these steps:

1. For participants who do not provide data, assume that they had no improvement to report. This is a conservative assumption. (Guiding Principle 6)
2. Check each value for realism, usability, and feasibility. Discard extreme values, omitting them from the analysis. (Guiding Principle 8)
3. Assume only first-year benefits, even though some programs will add value in the second and third year. (Guiding Principle 9)
4. Adjust the improvement from step 3 according to the confidence level provided by the participant, multiplying it by the confidence percentage. The confidence level represents the error in the estimate as suggested by the participant. For example, a participant indicating 80 percent confidence in the data provided is reflecting a 20 percent possibility of error. In a $10,000 estimate with an 80 percent confidence factor, the participant is suggesting that the value could be in the range of $8,000 to $12,000. To be conservative, the lower number is used. Thus, the amount of improvement is multiplied by the confidence factor. (Guiding Principle 7)
5. Adjust the new values according to the percentage of improvement related directly to the program, using straight multiplication. This calculation isolates the effects of the program from other influences. (Guiding Principle 5)

The monetary values that participants determine by following these five steps are combined to arrive at a total program benefit. Since the values are already annualized (assuming first-year benefits only), the total represents the annual benefit of the program. This value is placed in the numerator of the formula for calculating the ROI.

### Applying Action Plans

The impact of the action plan process can be impressive. For example, consider the following case. In a medium-sized manufacturing facility, a program that focused on improving interpersonal skills with employees was developed for first-level supervisors. Several of the areas addressed were productivity improvement, scrap reduction, absenteeism, turnover, grievances, and safety. These areas were discussed, and supervisors learned the skills to make improvements in each area. Supervisors were required to develop action plans for improvement and to report their results in a follow-up six months after the program. In this situation, the improvement measures were predetermined on the basis of the needs assessment. The following results were documented from a pilot group:

- The department unit hour was increased from 65 to 75. The unit hour is a basic measure of productivity; a unit hour of 60 is considered average and acceptable work.
- Scrap was reduced from 11 percent to 7.4 percent.
- Absenteeism was reduced from 7 percent to 3.25 percent.
- The annual turnover rate was drastically reduced from 30 percent to 5 percent.
- Grievances were reduced by 80 percent.
- Lost-time accidents were reduced by 95 percent.

These results were achieved by supervisors' practicing what they had learned and reporting the results of their action plans. Although these results are impressive, three additional steps are needed to develop the ultimate evaluation, the ROI. First, the amount of the improvement that is actually linked to the program must be determined, working with each measure. In this situation, supervisors estimated the percentage of the improvement that could be directly linked to the program. For example, while absenteeism showed an overall decrease of 3.75 percent, the supervisors collectively estimated that only 46 percent of the reduction was actually linked to the program. Therefore, the 3.75 percent reduction in absenteeism was reduced to 1.725 percent. This figure was then further adjusted by factoring in the participants' confidence levels. In this example, supervisors were 84 percent confident about their allocation of the absenteeism improvement. When adjusted for the 84 percent confidence level, 1.725 percent became 1.45 percent. These two adjustments isolated the effects of the program on the output measures; this process is fully described in the next book in this series, *Isolation of Results*.

The second step in developing the ROI measurement is converting the data to monetary values. In our example, a value for a single absence needed to be determined and used to calculate the annual benefit of the improvement in absenteeism. There are at least ten ways to place values on data, and they are fully described in the fourth book in this series, *Data Conversion*. In this case, the supervisors had developed an estimated value of one absence, which had been used in several previous applications that required a value for the cost of absenteeism. Consequently, the total number of absences avoided was calculated and multiplied by the value of one absence to obtain the training program's annual impact through reduction of absenteeism. This process shows the economic value of the program on that specific output measure. The two steps of isolating the effects of the program and converting data to monetary values were performed for each of the six improvement measures;

after the total annual value of each of the six improvements was calculated, the total value of all six represented the annual economic benefit of the program.

The third step necessary to calculate ROI is developing the fully loaded costs of the program. In this step, the costs related to needs assessment and program development are prorated based on the number of people involved or the number of times a program is implemented over the lifetime of the program. In addition, all direct program costs are captured, along with the cost of the participants' salaries and benefits during the time that they are involved. The fully loaded cost for all participants reflects the total investment in the program for this group. The process of arriving at the total cost figure is fully explained in the fifth book in this series, *Costs and ROI*.

After the preceding three additional steps have been completed, ROI can be calculated, using the formulas described in *ROI Fundamentals* (net benefits divided by costs multiplied by 100). In our example, after all six improvement items were converted to monetary units, the annual benefits directly attributed to the program totaled $775,000. The fully loaded costs for the program—which included the costs of needs assessment and program development, delivery, and evaluation—came to $65,000. Thus, the ROI was calculated as follows:

$$\text{ROI} = \frac{\text{Net Program Benefits}}{\text{Program Costs}} \times 100$$

$$\text{ROI} = \frac{\$775{,}000 - \$65{,}000}{\$65{,}000} \times 100 = 1092\%$$

This impressive ROI has credibility because the adjustments made to the data make it a conservative estimate. Without these three steps, the target audience might have wondered what part of the results was actually linked to the program and whether the benefits actually exceeded the costs.

### Identifying Advantages and Disadvantages of Action Plans

Although action plans have many advantages, their use raises at least two concerns. First, the process relies on direct input from participants, usually with no assurance of anonymity. Therefore, the information may be biased and unreliable. Also, action plans can be time-consuming for participants; as a result, if a participant's supervisor is not active in the process, the participant may not complete the assignment.

As this section has illustrated, the use of action plans offers many advantages. Action plans are simple and easy to administer, easily understood by participants, appropriate for a wide variety of programs, and appropriate for collecting all types of data. In addition, they can be used to measure reaction, learning, application, and business impact. They can be used with or without other evaluation methods. The two disadvantages can be overcome with careful planning and implementation. Building action plans into the program mitigates the presence of an evaluation. It suggests, instead, that follow-through on actions that drive business performance is an important part of the program itself. It also provides participants the opportunity to track how successful they are with their implementation. When participants are prepared in advance, their apprehension about participating is reduced. Ensuring participants that their results will be used for program or process improvement provides some comfort that their individual job performance is not being examined. Because of their tremendous versatility and the adjustments that can be made during analysis to render more conservative estimates, action plans have become an important data collection tool for ROI evaluations.

## Performance Contracts

Performance contracts are a slight variation of action plans, with the added feature of a pre-program commitment. Based on the principle of mutual goal setting, a performance contract is a written

agreement between a participant and the participant's manager. The participant agrees to improve performance in an area of mutual concern related to the content of the program. The agreement takes the form of a project to be completed or a goal to be accomplished soon after the program's completion. The agreement spells out what should be accomplished, at what time, and with what results.

Performance contracting is administered in much the same way as the action planning process. Although the steps depend on the specific kind of contract and on the organization, a common sequence of events is as follows:

1. With his or her manager's approval, the participant decides to be involved in a program or project.
2. The participant and his or her manager agree on a topic for improvement and specific measures of that improvement.
3. Specific, measurable goals are set.
4. The participant undergoes the program, during which the contract is discussed and plans are developed to accomplish the determined goals.
5. Following the program, the participant works on the contract, adhering to a specific deadline.
6. The participant reports the results to his or her immediate manager.
7. The manager and participant document the results and forward a copy of the documentation to the evaluation team, along with appropriate comments.

Participants and their managers together select the area or measure to be improved prior to the start of the program. The process of selecting the area for improvement is similar to the process of developing an action plan, described earlier in this chapter. The

proposed improvement can cover one or more of the following areas:

- *Routine performance,* which includes specific improvements in routine performance measures such as production targets, efficiency, or error rates
- *Problem solving,* which focuses on specific problems such as an unexpected increase in accidents, a decrease in efficiency, or a loss of morale
- *Innovative or creative applications,* which include initiating changes or improvements in work practices, methods, procedures, techniques, or processes
- *Personal development,* which involves learning new information or acquiring new skills in order to increase individual effectiveness

The proposed improvement should be stated in terms of one or more objectives. The objectives should state what will be accomplished when the contract is complete. These objectives should be

- Written
- Understandable (by all involved)
- Challenging (requiring unusual effort to achieve)
- Achievable (something that can be accomplished)
- Largely under the control of the participant
- Measurable
- Time sensitive

The detailed steps needed to accomplish the contract objectives are developed by following the guidelines for developing action plans items, which were presented earlier in this chapter. The methods for analyzing the data and reporting progress are also essentially the same as those used in the action planning process.

## Final Thoughts

This chapter has provided an overview of three data collection approaches that can be used in ROI evaluation. Use of these methods in collecting data for ROI calculations is gaining more acceptance. Performance monitoring, action plans, and performance contracting are often used to collect data for ROI evaluations.

### References

Mondschein, M. *Measurit: Achieving Profitable Training*. Leawood, Kans.: Leathers, 1999.

Phillips, J. J., and Phillips, P. P. "Evaluating the Impact of a Graduate Program in a Federal Agency." In P. P. Phillips and J. J. Phillips (eds.), *In Action: Measuring ROI in the Public Sector*. Alexandria, Va.: ASTD, 2002.

# 5

# Measuring Reaction and Planned Action

This chapter focuses on measurement of reaction and planned action (Level 1 in the ROI Methodology), which involves collecting data at the beginning of and during a program. Participant feedback supplies powerful information that can be used to make adjustments and measure success. This chapter outlines the most common approaches to collecting Level 1 data and explores ways to use the information for maximum value.

## Why Measure Reaction and Planned Action?

It is difficult to imagine a program being conducted without the collection of feedback from those involved in the program or, at least, from the participants. Collecting reaction data serves several purposes. Participant feedback is critical to understanding how well a program serves the customer and the potential of the program to meet identified business needs.

### Customer Satisfaction

Reaction is a measure of customer satisfaction with a program. Without sustained favorable reactions, a program may not succeed. The individuals who have a direct role in planning and implementing a program are immediately affected by reaction data and often have to change processes or procedures or make other adjustments

in response. Participant feedback on preferences is critical to making adjustments and changes in a program as it unfolds. Feedback from program supporters is also important because this group is in a position to influence the program's continuation and development. The sponsors—who approve budgets, allocate resources, and ultimately, live with the program's success or failure—must be completely satisfied with the program, and their overall satisfaction must be verified early and often.

### Immediate Adjustments

A program can go astray quickly, and sometimes a program ends up being the wrong solution for the specified problem. A program can also be mismatched to the problem from the beginning, so getting feedback early in the process is necessary so that immediate adjustments can be made. This can help prevent misunderstandings, miscommunications, and, more important, misappropriations. Gathering and using reaction data promptly can allow an improperly designed program to be changed before more serious problems arise.

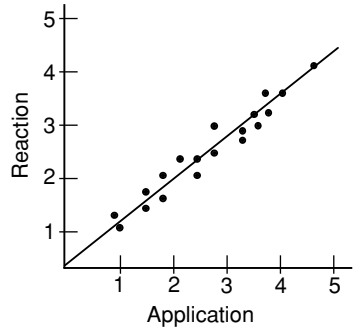
### Team Evaluation

In some projects and programs, reaction or feedback data are used to evaluate the effectiveness of the team that has implemented the program. In learning and development, consulting, and change management, facilitator evaluation is critical. These data can reveal how well facilitators have fulfilled their responsibilities. Reaction data may also be used to evaluate other team members, including program coordinators, program planners, or sponsors. Such input also sometimes leads the program team to make quick adjustments in order to ensure that the program progresses properly.

### Predictive Capability

A relatively recent use of reaction data is predicting the success of a program through analytical techniques. Participants are asked to predict the effectiveness of the application of learning from the

**Figure 5.1. Correlations Between Reaction and Application Data**




program and, in some cases, the resulting value of that application. The reaction data thus become a forecast. Figure 5.1 shows the correlation between reactive feedback and application data. Studies have been conducted that verify this correlation.

In this type of analysis, the reaction measures are taken as the program is introduced, and the success of the application (or implementation) is later judged using the same scale (a rating from 1 to 5). When significant positive correlations between reaction and application measures are present, reaction measures can have predictive capability. Some reaction measurement items shown to have predictive capability are

- The program is relevant to my job.
- The program is necessary.
- The program is important to my success.
- I intend to implement the program.
- I would recommend that others pursue similar programs.

Measurement items such as these consistently lead to strong positive correlations and consequently represent more powerful feedback than typical measures of satisfaction with a program. Some

organizations collect these or similar reaction measures for every project or program initiated.

## Importance of Other Levels of Evaluation

Feedback data are critical to a program's success and should be collected for every program. Unfortunately, however, in some organizations, feedback alone has been used to measure program success. In the meetings and events industry, which is a $150 billion industry, it is estimated that 95 percent of meetings and events are evaluated only at the reaction level (Phillips, Myhill, and McDonough, 2007). For example, in a financial services firm in Israel, the traditional method of measuring the effectiveness of an ethics program relied entirely on feedback data from employees, asking them whether the ethics policy was appropriate, fair, and necessary. Positive feedback is obviously critical to the policy's acceptance but is no guarantee that the new policy will be successfully executed. As subsequent policy changes were made, executives became interested in more realistic evaluation, which included the extent to which employees actually understood the policy (learning), the extent to which employees followed the policy in their work (application), and the effectiveness of the policy in reducing ethical violations and infractions (impact). Only when these additional measures were taken could the full scope of success be identified.

Collecting only reaction data for a project or program is risky. Research shows that a positive reaction to a program does not always correlate to positive program results Although reaction data are easy to collect, they are based on many elements and contain potential bias. Data collected later in the evaluation chain are more helpful for evaluating projects or programs.

## Areas of Feedback

In capturing reaction and planned action data, it is important to focus on the content of the project, program, or initiative. Too often,

**Table 5.1. Noncontent Issues Versus Content Issues for a Marketing Event**

| Noncontent Issues | Content Issues |
|---|---|
| Demographics | Service quality |
| Facilities | Learning environment |
| Location | Relevance of program to participant's clients |
| Transportation | |
| Registration | Importance of program to participant's success |
| Logistics | |
| Hotel service | Timing of program |
| Media | Use of participant's time |
| Food | Amount of new information |
| Breaks and refreshments | Quality of presentations |
| Cocktail reception | Usefulness of materials |
| Gala party | Perceived value of program |
| Closing dinner | Value of contacts made |
| Opening keynote | Planned use of material |
| Quality of speakers | Forecast of impact on sales |
| Future needs | Confidence of estimate of impact |
| Overall satisfaction | Overall satisfaction with program content |

feedback data reflect aesthetic issues that are not relevant to the substance of the project. Table 5.1 shows two types of reaction data that could be captured at a marketing event for customer relationship managers by using a reaction questionnaire. The difference is more than subtle. The traditional way to evaluate activities is to focus on noncontent issues or inputs. In Table 5.1, the column on the left represents areas important to the activity surrounding the marketing event but contains nothing indicating results or possible results achieved from the event. The column on the right reflects a focus on content. This is not to suggest that the nature of the service, the atmosphere of the event, and the quality of the logistics are not important; it is assumed that these issues will be taken care

of and addressed appropriately. However, a more important set of data, focused on results, incorporates detailed information about the perceived value of the meeting, the importance of the content, and the planned use of material or a forecast of the impact—indicators that successful results were achieved.

Many topics are targets for reaction feedback. Feedback data are needed in connection with almost every major issue, step, or process in order to make sure that things are working properly. The areas of feedback used for reaction data depend, to a large extent, on the organization and the purpose of the evaluation. Some requests for feedback are simple, while others are detailed and require more time for participants to answer and for evaluators to collect. Feedback requests should be designed to supply the information needed to satisfy the purpose of the evaluation. Following is a comprehensive list of over twenty common types of feedback:

- *Progress on objectives*. To what degree were the objectives met?
- *Program content*. Was the content appropriate? Important? Motivational?
- *Program materials*. Were the materials useful? Practical?
- *Pre-work materials*. Were the pre-work, or introductory, materials necessary? Helpful?
- *Assignments*. Were the assignments helpful?
- *Method of delivery*. Was the method of delivery appropriate? Efficient? Timely?
- *Facilitator*. Was the facilitator effective? Knowledgeable? Responsive?
- *New information*. How much new information was included?

- *Motivation to learn*. Were you motivated to learn the content?
- *Relevance*. Was the program relevant to your needs?
- *Importance*. How important was the content to your success?
- *Value of program*. Was this a good investment? A good use of your time?
- *Rewarding*. Was the program rewarding?
- *Challenging*. Was the program difficult?
- *Logistics*. Were the scheduling and coordination efficient?
- *Facilities*. Did the facilities enhance the program? The learning environment?
- *Potential barriers*. What are the potential barriers to application of the program material?
- *Use of material*. How will you apply what you have learned?
- *Planned improvements*. What improvements will you make?
- *Recommendations for target audiences*. What is the appropriate audience for this program? Will you recommend it to others?
- *Overall evaluation*. What is your overall rating of the program?

Objective questions covering each of the areas listed will ensure thorough feedback from participants; although, if the data are not going to be used to improve the program or provide relevant

information, then a type may be eliminated. Such feedback can be extremely useful in making adjustments in a program and may assist in predicting performance after the program.

In training, learning, development, consulting, and change management programs, evaluation of the facilitator deserves additional comment. In some organizations, the primary evaluation centers on the facilitator (a separate form may be used for each facilitator, if there are several), covering a variety of areas such as the following:

- Preparation for sessions
- Knowledge of the subject matter, including familiarity with content and depth of understanding
- Presentation skills, including clarity of the presentation, pacing of material, and eye contact
- Communication skills, including use of understandable language and real-life examples, as well as promotion of discussion
- Assessment of learners' understanding, and responses appropriate to learners' needs and questions
- Use of appropriate technology and ability to respond effectively to technical requirements of learners
- Encouragement of application of learning through the use of real-life examples, job-related discussions, and relevant exercises

In organizations in which significant learning activity occurs, reaction data collection is usually automated, employing computerized scanning and reporting. Typical questions can easily be developed for a scan sheet, and reports that will help decision makers understand and use the data can be programmed. Some

organizations use direct input into a Web site to develop not only detailed reports but also databases, allowing feedback data to be compared with data from other programs or from the same program with other facilitators.

## Data Collection Issues

Several issues can affect the quality and quantity of data collected at Level 1. These issues should be addressed professionally and efficiently in order to ensure that data are sufficient in quality, quantity, and objectivity.

### Timing

The timing of data collection centers on particular events connected with the program. For example, reaction data collection for any industry conference may occur at the end of every session as well as at the end of the conference itself. Reaction data may be collected immediately after content or information is presented and then again after participants have had time to put that content into action. As discussed previously, feedback during the early stages of program implementation can be extremely useful. Ideally, early feedback will validate the decision to go forward with the project and confirm that the project aligns with business needs. Noting problems in the initial feedback allows adjustments to be made early in the program's implementation. In practice, however, many organizations omit early feedback and wait until significant parts of the project have been implemented, when feedback may be more meaningful.

Longer projects may require data collection at multiple points, which may require increased coordination and project management than when data are collected only once. However, automation makes data collection at multiple points in time more manageable. Measures can be taken at the beginning of the project and then at routine intervals once the project is under way.

### Methods

A variety of methods can be used to collect reaction data. Instruments range from simple surveys to comprehensive interviews. The appropriate method depends on the type of data needed (quantitative or qualitative), the convenience of the method for potential respondents, the culture of the organization, and cost.

Questionnaires or surveys are the most common method of collecting and measuring reaction data. Questionnaires and surveys come in all sizes, ranging from short forms to detailed, multiple-page instruments. They can be used to obtain subjective data about participants' reactions as well as to document responses for use in a projected ROI analysis. Here participants forecast improvement in business measures as well as attempt to identify a monetary contribution from those improvements. Proper design of questionnaires and surveys is presented in Chapter One.

Interviews, though not used as frequently as questionnaires to capture reaction data, may be conducted to secure data that are difficult to obtain through written responses. Interviews may be desirable when a program is new or has a long duration. They can uncover success stories that may help to communicate early achievements of the program. Respondents may be reluctant to describe their experiences using a questionnaire but may volunteer the information to a skillful interviewer using probing techniques. A major disadvantage of interviews is that they are time-consuming, which increases the cost of data collection. They also require interviewer preparation in order to ensure that the process is consistent. Thus, for most reaction data collection, interviews are too expensive.

As previously discussed, focus groups are a useful tool when in-depth feedback is needed. Focus groups can be used in lieu of or in addition to the standard end-of-program questionnaire. In lieu of the questionnaire, the evaluator will gather limited data, and those data will be focused on one or two key issues in the program content and delivery. When using a focus group in addition to a questionnaire, specific follow-up on the initial results can

occur. For example, a large-scale performance improvement initiative took place in a federal government agency. The reaction data from the end-of-program questionnaire indicated that even though participants were satisfied overall, they saw neither the relevance nor the value of the program itself. Focus groups allowed participants to express their concerns with the program more adequately than they could through the questionnaire.

### Administrative Guidelines

Several administrative guidelines can improve the effectiveness of data collected at Level 1.

*Keep responses anonymous.* Anonymous feedback is recommended. It allows participants to be open with their comments, resulting in helpful and constructive feedback. Otherwise, the input may be biased or stifled because of concerns about the facilitator's reaction.

*Have a neutral person collect the data.* In addition to anonymous responses, it is helpful to have a neutral person to collect the feedback data. In some organizations, the program coordinator or sponsor conducts the evaluation at the end of the program, independent of the facilitator. This action increases the objectivity of the input and decreases the likelihood of the facilitator reacting unfavorably to criticism contained in the feedback.

*Provide questions in advance.* For lengthy evaluations involving programs that span several days, distribute the feedback questions early in the program so that participants can familiarize themselves with the issues. Participants can address specific topics as they are discussed during the program and will have more time to think through particular issues. They should be cautioned, however, not to reach a final conclusion on general issues until the end of the program or project.

*Explain the purpose of the feedback and how it will be used.* Although this is sometimes understood, repeating where the information goes and how it is used in the organization can be helpful. There may still

be some mystery surrounding the use of feedback data. Restating the process in terms of the flow of data and the use of data may clarify this issue.

*Consider ongoing evaluation.* In lengthy programs, an end-of-program evaluation may suffer because participants are unable to remember their reactions from earlier in the program. An ongoing evaluation can be used to improve this situation. One approach is to distribute evaluation forms at the beginning of the program and then explain when and how to supply the information. After each topic is presented, participants evaluate that topic. That way, the information can be easily recalled by participants, and the feedback is more useful to program evaluators. Another approach is to use a daily or routine feedback form to collect input on program pacing, degree of involvement, unclear items, and so on. Exhibit 5.1 shows a routine feedback form.

*Collect information related to improvement.* Although it is difficult to secure realistic input related to cost reductions or savings as part of feedback data, it is worth a try. The response may be surprising. At times, a simple question will cause participants to focus on savings or improvements. A statement to elicit this sort of data might look like this:

> Please estimate the savings in monetary values that will be realized (for example, from increased productivity, improved methods, or reduced costs) over a period of one year as a result of this program: $ _____
>
> Please explain the basis of your estimate.
>
> Express as a percentage the confidence you place on your estimate (0% = no confidence, 100% = certainty): _____

*Allow ample time for participants to provide data.* A time crunch can cause problems. If participants are asked to provide feedback in a hurry, they may provide incomplete information, cutting their feedback short in an effort to finish and leave. One way to avoid

**Exhibit 5.1. Routine Feedback Form**

1. What issues still remain confusing or unclear?
   _____
   _____
   _____
2. The most useful information was
   _____
   _____
   _____
3. It would help me if you would
   _____
   _____
   _____
4. The pacing of the program was
   - ☐ Just right
   - ☐ Too slow
   - ☐ Too fast
5. Three important items that you should cover soon are
   1) _____
   2) _____
   3) _____
6. Comments
   _____
   _____
   _____

this problem is to allow ample time for evaluation in a scheduled session before the end of the program. The evaluation session could be followed by a wrap-up of the program. A thirty-minute session would provide an opportunity for thorough feedback, enhancing the quality and quantity of information.

## Uses of Reaction Data

Sometimes, feedback is solicited, tabulated, summarized, and then disregarded. Too often, program evaluators use the material to feed their egos and then let it quietly disappear in the files, forgetting the

original purposes for its collection. Information should be collected and used for one or more of the purposes of evaluation; otherwise, the exercise is a waste of time. Typical purposes of collecting reaction data include these:

*Monitoring customer satisfaction.* Reaction data indicate participants' and other stakeholders' overall reaction to and satisfaction with the program, allowing program developers and owners some insight into how satisfied the customers are with their product.

*Identifying strengths and weaknesses of a program.* Feedback helps identify the weaknesses as well as the strengths of a program. Finding weaknesses often leads to adjustments and changes as a result of the feedback. Identifying strengths so that they can be replicated in future designs is an important use of reaction data.

*Developing norms and standards.* Because reaction data are usually collected for 100 percent of programs and because tabulation of the data can be automated, it is relatively easy to develop norms and standards for reaction data within an organization. Target ratings can be set, and later, the actual ratings can be compared against those norms and standards.

*Evaluating facilitators and team members.* Perhaps one of the most common uses of reaction data is program team evaluation. If data collection is properly planned and executed, helpful feedback can be provided to facilitators and program leaders, allowing them to make adjustments that increase their program's effectiveness. Some caution needs to be exercised because facilitator and program leader evaluations are sometimes biased due to the presence of a facilitator during the evaluation or to relationships with team leaders and members. When a concern with bias does exist, other evidence may be necessary to provide a complete performance assessment.

*Evaluating planned improvements.* Feedback data from program participants can provide a profile of their planned improvements. This profile can be compared with the on-the-job actions that were anticipated results of the program when it began. Thus, reaction data can be a rich source of information on what processes, methods,

or behaviors participants may be changing or implementing because of what they have learned.

*Linking with follow-up data.* If a follow-up evaluation is planned, linking reaction data with follow-up data may be helpful in order to see whether planned improvements became reality. In most cases, planned actions are inhibited in some way by on-the-job barriers.

*Obtaining material for marketing programs.* For some organizations, participant feedback data provides helpful marketing information. Direct participant statements or reactions on particular issues can provide information that may be convincing to potential participants. Program brochures and other marketing pieces often contain quotes from and summaries of feedback data.

## Final Thoughts

This chapter discussed data collection at the first level of evaluation, reaction and planned action. Measuring reaction is a component of every study and is a critical factor in a program's success. Data are collected by means of a variety of techniques, although surveys and questionnaires are most often used, due to their cost-effectiveness and convenience. Level 1 data are important because they allow immediate adjustments to be made to a program. While reaction data are important to those who are directly involved in implementing a program, their value to executives is usually low. The value of data to executives increases as the evaluation moves up the chain of impact.

### Reference

Phillips, J. J., Myhill, M., and McDonough, J. B. *Proving the Value of Meetings and Events.* Birmingham, Ala.: ROI Institute; Meetings Professional International, 2007.

# 6

# Measuring Learning and Confidence

Measuring learning is an important part of the evaluation process, especially when a program is intended to change on-the-job behaviors or processes. Participants' knowledge of what to do and how to do it is critical to a program's success. This chapter focuses on simple, commonly used techniques for measuring learning and begins with a look at the reasons for measuring learning.

## Why Measure Learning and Confidence?

Several key principles illustrate the importance of measuring learning during the course of a program. Each in itself is sufficient to justify measurement of learning; together, they provide an indication of the full range of benefits that result from measuring the changes in skills, knowledge, and other qualities that occur during a program.

### The Learning Organization

During the past two decades, organizations have experienced a rapid transformation of competitive global markets as a result of economic changes. Organizations need new ways to serve customers, and they need new technology and innovations to enhance efficiency, to restructure, to reorganize, and to execute their functions

globally. In response to this need for strategic change, the concept known as the learning organization evolved. This concept requires organizations to use learning in proactive and integrated ways to support growth for individuals, teams, and entire organizations. Peter Senge popularized the idea of the learning organization, suggesting that organizations capture, share, and use knowledge so that their members can work together to change how the organization responds to challenges (Senge, 1990). Managers must question old social constructs and practice new ways of thinking.

Learning must take place within and must support a framework of teams and larger groups in which individuals can work together to generate new knowledge. The process must be continual, because a learning organization is a never-ending journey (Watkins and Marsick, 1996).

With the new focus on creating learning organizations in which countless activities and processes are in place to promote continual learning, measurement of learning has become an important issue. How can we know whether an organization has become a learning organization? How is learning in such an organization measured? Can learning be measured on a large scale?

### Compliance Issues

Organizations face an increasing number of regulations with which they must routinely comply. These regulations involve all aspects of business, and governing bodies consider them essential to protect customers, investors, and the environment. Employees must have some knowledge of the regulations in order to remain in compliance. Therefore, an organization must measure the extent of employee learning and understanding of the regulations to ensure that compliance does not become a problem.

Some programs are compliance-driven. For example, one large banking organization had to implement a major program to ensure that its employees were all familiar with regulations on money laundering. This program was precipitated by the bank's continuing

failure to comply with the regulations. The problem appeared to arise from a lack of knowledge of the rules. When programs such as this are initiated, learning must be measured.

### Development of Competencies

The use of competencies and competency models has increased dramatically in recent years. In the struggle for competitive advantage, many organizations have focused on people as the key to their success. Competency models are used to ensure that employees do the right things, clarifying and articulating what is required for effective performance. Competency models help organizations align behavior and skills with the strategic direction of the company. A competency model describes a particular combination of knowledge, skills, and characteristics necessary to perform a role in an organization. Competency models are used as tools for recruiting, selecting, training, reviewing performance, and even removing individuals from the organization (Lucia and Lepsinger, 1999). With the increased focus on competencies, measuring learning is a necessity.

### Certification

The learning and development field has gone crazy over certification. Employees are becoming certified in many processes, ranging from Microsoft™ products and service to Six Sigma implementation. With this focus on certification comes an increased emphasis on measuring learning. Almost every certification program (if not all of them) requires that individuals demonstrate that they know something or know how to do something. Certification programs place much more emphasis on measurement than typical learning and development programs.

### Consequences of an Unprepared Workforce

Perhaps the most important reason to focus on learning measurement is to ensure that the workforce is prepared. Many sad and

disappointing stories have detailed how employees are not capable of performing the skills needed to do their jobs and deliver excellent customer service. Hearing about or experiencing such gaps in workers' skills has motivated many management teams to work hard to ensure that their workers are prepared. The only way to make sure that employees have learned the knowledge and skills needed to perform their jobs successfully is to measure learning through some credible, valid, and reliable process.

### The Role of Learning in Programs

Although some programs involve new equipment, processes, and technology, the human factor remains critical to the success of most programs. Whether an organization is restructuring or adding new systems, employees must learn how to work in the new environment, a process that requires them to develop new knowledge and skills. Simple tasks and procedures do not necessarily come with new processes. Instead, complex environments, procedures, and tools must be used in an intelligent way to reap the desired benefits for the organization. Employees must learn in different ways—not just in a formal classroom environment but through technology-based learning and on-the-job practice. Team leaders and managers serve as coaches or mentors in some programs. In a few cases, learning coaches or on-the-job trainers are used to ensure that learning is transferred to the job and is implemented as planned.

Participants don't always fully understand what they must do. Although the chain of impact can be broken at any level, a common place for such a break is at Level 2, learning and confidence. Employees simply may not know what to do or how to do it properly. When application and implementation of program skills do not go smoothly, program leaders need to determine whether a learning deficiency is the problem; if it is, they may be able to eliminate it. In other words, measurement of learning is necessary to

leaders' understanding of why employees are or are not performing as they should.

## Measurement Issues

Several factors affect the nature and scope of measurement at the learning level. These factors include challenges of testing, program objectives, the measures themselves, and timing.

### Challenges

The greatest challenge when measuring learning is maintaining objectivity without crossing ethical or legal lines while keeping costs low. A common method of measuring learning is testing (discussed in Chapter Two). This approach generates its own unique challenges.

The first challenge is the fear factor. Few people enjoy being tested. Many are offended by it and feel that their professional expertise is being questioned. Some people are intimidated by tests, which bring back memories of their third-grade math teacher, red pen in hand.

Another challenge raised by tests is the legal and ethical repercussions of basing decisions about employees' job status on test scores. Therefore, organizations often use other techniques to measure learning, such as surveys, questionnaires, role-playing, and simulations. These methods, however, bring their own challenges as well—most notably, the financial burden they impose and the potential for inaccurate measures. Trade-offs often must be made between resources and the accuracy of the learning measurement process.

### Program Objectives

The starting point for any level of measurement is development of the program's objectives. Measurement of learning is built on a

program's learning objectives. Learning and development professionals are skilled at generating detailed learning objectives by following the process described in book one of this series, *ROI Fundamentals*. However, even for programs in which the focus is not necessarily on a learning activity but on, for example, implementing a new policy, initiating a new procedure, or creating a wellness and fitness center, the first step is ensuring that learning objectives are in place. Following are some examples of learning objectives:

- Identify the six features of the new ethics policy.
- Demonstrate the use of each software routine within the standard time.
- Score 75 or better on the new-product quiz.
- Explain the value of diversity in a work group.
- Successfully complete the leadership simulation.
- Know how to apply for housing assistance.

Typically, the objectives of a program are broad and indicate only major skills or general knowledge areas that should be achieved as the program is implemented. These are sometimes called a program's key learning objectives. Key learning objectives can be broken into specific measures that define the criteria for success or failure. Such detail is necessary when a tremendous number of tasks, procedures, or new skills must be learned in order to make a program successful. For other programs, this level of detail might not be needed; identifying the major objectives and indicating what must be accomplished in order to meet each objective is often sufficient.

### Typical Measures

Measurement of learning focuses on knowledge, skills, and attitudes as well as on an individual's confidence in applying or implementing

the project or program as desired. Typical measures at this level involve

- Skills
- Knowledge
- Awareness
- Understanding
- Contacts
- Attitudes
- Capacity
- Readiness
- Confidence

Obviously, the more detailed the knowledge area, the greater the number of possible objectives. The concept of knowledge is quite broad and often includes the assimilation of facts, figures, and ideas. Instead of knowledge, terms such as *awareness*, *understanding*, and *information* may be used to identify specific categories of knowledge. Sometimes, a participant's perceptions or attitudes change based on what he or she has learned. For example, participants' perceptions of a diverse work group are often changed when a major diversity program is implemented. In some cases, the focus of a program is development of a reservoir of knowledge and related skills in order to improve capability, capacity, or readiness. Networking is often part of a program; in that case, developing contacts who may be valuable later is important. Networking may occur within or outside of an organization. For example, within an organization, a program may include people from different functional areas of the organization, and an expected outcome from a learning perspective might be knowing who to contact at particular times in the future. In programs that involve different organizations, such as a marketing event, new contacts that result from the event may be important and may ultimately pay off in terms of efficiency or revenue growth.

**Timing**

The timing of learning measurement can vary depending on the program content and design as well as the participants involved

in the program. In some situations, a preliminary measure needs to be taken, so a pretest is generated to determine the extent to which participants understand the specific objectives or content of the program. A pretest can be important in assessing participants' current skills and knowledge so that the learning of additional skills and knowledge can be planned more efficiently. This step may prevent participants from being taught information they already know. When a pretest has been used, it is common to administer a posttest in order to obtain data to compare with those from the pretest. The posttest can be administered early in the program or as soon as the learning portion is completed. The pretest and posttest should be conducted under the same or similar conditions, using questions or other test items that are identical or very similar. A factor to consider in using pretests and posttests is participants' existing level of knowledge. Participants who have no previous knowledge of the content presented should not be pretested. The baseline for them is zero. When they are pretested, the results may reflect what participants think they know, and the corresponding posttest may result in lower scores, because through the program process participants have realized what they don't know. Participants who are expected to have some level of knowledge are good targets for pretesting as the program is intended to enhance their current level of knowledge.

If no pretest is administered, measurement of learning can occur at various times. If formal learning sessions connected with the program are offered, a measure will be taken at the end of each session to ensure that participants are ready to apply their newly acquired knowledge. If a program has no formal learning sessions, measurement may occur at different intervals. In long-term programs, as skills and knowledge grow, routine assessment may be necessary to measure both the acquisition of additional skills and the retention of previously acquired skills. The timing and frequency of measurement are aligned with the need to know the new information and are always balanced against the cost of obtaining, analyzing,

and responding to the data. Ideally, the timing of measurement is considered as part of the development of a data collection plan, as discussed in Chapter Six of *ROI Fundamentals*, the first book in this series.

**Data Collection Methods**

One of the most important considerations in measuring learning is the specific way in which data are collected. Learning data can be collected using many different methods. The following list of instruments includes just some of the data collection methods that are used:

- Questionnaires
- Performance tests
- Technology simulations and task simulations
- Case studies
- Role-playing or skill practice
- Informal assessments

These methods were discussed in earlier chapters.

Many times, learning will be measured by means of a self-assessment instrument on which participants indicate the extent to which they have learned the material. Self-assessments were discussed briefly in Chapter Two. They are discussed here because of their frequent use in measuring learning. Exhibit 6.1 shows an end-of-the-conference questionnaire for an annual business development conference for insurance agents. The questionnaire includes questions that cover reaction and learning, a typical data collection method that provides an opportunity to raise the percentage of programs that are evaluated at Level 2. In essence, the learning measures are taken at the same time as the reaction measures; in

Exhibit 6.1. End-of-Conference Questionnaire

| Reaction | Strongly Disagree 1 | 2 | 3 | 4 | Strongly Agree 5 | n/a |
|---|---|---|---|---|---|---|
| 1. The meeting was organized and efficient. | ○ | ○ | ○ | ○ | ○ | ○ |
| 2. The speakers were effective. | ○ | ○ | ○ | ○ | ○ | ○ |
| 3. The meeting was valuable for business development. | ○ | ○ | ○ | ○ | ○ | ○ |
| 4. The meeting content was important to my success. | ○ | ○ | ○ | ○ | ○ | ○ |
| 5. The meeting was motivating. | ○ | ○ | ○ | ○ | ○ | ○ |
| 6. The meeting was challenging. | ○ | ○ | ○ | ○ | ○ | ○ |
| 7. The meeting contained new information. | ○ | ○ | ○ | ○ | ○ | ○ |
| 8. The meeting represented an excellent use of my time. | ○ | ○ | ○ | ○ | ○ | ○ |
| 9. I will use the material from this conference. | ○ | ○ | ○ | ○ | ○ | ○ |
| **Learning** | | | | | | |
| 10. I can identify the five steps for a business development strategy. | ○ | ○ | ○ | ○ | ○ | ○ |
| 11. I can develop a business development plan. | ○ | ○ | ○ | ○ | ○ | ○ |
| 12. I can select the best community service group to join for business development. | ○ | ○ | ○ | ○ | ○ | ○ |
| 13. I can explain the changes in products. | ○ | ○ | ○ | ○ | ○ | ○ |
| 14. I can identify the five most effective ways to turn a contact into a sale. | ○ | ○ | ○ | ○ | ○ | ○ |
| 15. I can identify at least five agents to call for suggestions and advice. | ○ | ○ | ○ | ○ | ○ | ○ |
| **Suggestions for Improvement** | | | | | | |
| 16. What suggestions do you have for improving this conference? | ○ | ○ | ○ | ○ | ○ | ○ |

this instance, as mentioned, the data are collected at the end of the conference.

Another common method of measuring learning is the true-false test. This basic form of objective learning assessment tests recall of key concepts and information presented during a program. Answers are either right or wrong, unless the answers are intended to generate discussion rather than to actually test knowledge. If this is the case, it may be more appropriate to refer to this type of assessment as a true-false survey, as testing rules have probably not been followed.

## Administrative Issues

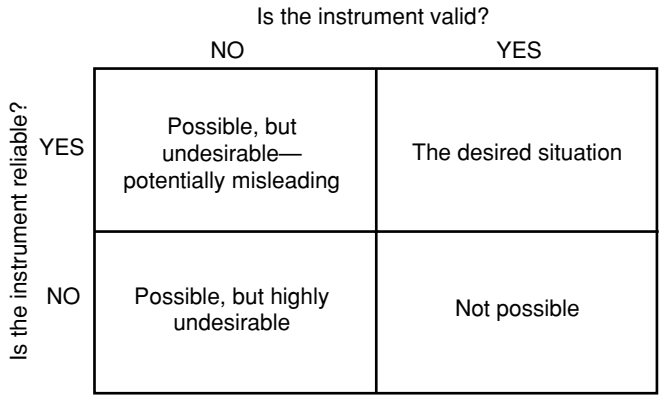
Several administrative issues must be addressed when learning is measured. Each issue is briefly discussed in this section and should be considered as part of the overall plan for collecting learning data.

### Validity and Reliability

Two important issues in test design are validity and reliability. Validity is the extent to which an instrument measures what it is designed to measure. Reliability is the extent to which the results provided by an instrument are stable or consistent over time. Any instrument used to collect data should be both valid (measure what it should measure) and reliable (provide consistent results over time). An instrument is reliable when the same questions, asked at different times, with no intervening processes or variables changing in the respondent's knowledge, yield the same responses. Significant deviations indicate that an instrument is unreliable.

Validity and reliability become particularly important criteria when a human resource action (job status change) is taken as a result of a person's failing a specific test or failing to meet a specific standard when learning is measured. For example, if an individual is promoted, denied promotion, provided an increase in pay, or assigned a job because of his or her performance on a test, the

**Figure 6.1. Relationship Between Validity and Reliability**




instrument must be defensible. Of course, in the vast majority of programs, the consequences of not passing a test will not be so severe. The concepts of validity and reliability and how to check for adequate levels of the two are beyond the scope of this book, but other sources provide more detail (for example, see Phillips and Phillips, 2007).

Understanding the relationship between validity and reliability is important (see Figure 6.1). An instrument cannot be valid unless it is also reliable, although, it can be reliable without being valid. Validity should be strived for when developing data collection instruments.

### Consistency

Tests, exercises, and assessments for measuring learning must be administered consistently from one group to another in order to effectively measure and compare learning between groups. Consistency refers to the time allotted to respond, the conditions under which participants complete the process, the resources available to them, and the amount of assistance they receive from other members of the group. These concerns can easily be addressed in the instructions.

When formal testing is used, participants should be monitored as they complete the test. Monitoring ensures that individuals work independently and that someone is available to provide assistance or answer questions as needed. These considerations may not apply in all situations, but they should be addressed in the evaluation plan.

### Pilot Testing

It is advisable to test an instrument with a small group to ensure that the instrument is both valid and reliable. A pilot test provides an opportunity to resolve anything that is confusing about the instructions, questions, or statements. When a pilot test is administered, it should be timed to see how long individuals take to complete it. Also, the individuals taking the pilot test should provide input on other ways to ask the questions, improve the flow of information, and generally improve the test. At a minimum, a test should be pilot-tested in order to examine its content. All too often, a test or survey is administered that does not cover the content necessary to support implementation of the program.

### Scoring and Reporting

Scoring instructions for the measurement process need to be developed in such a way that the person evaluating the responses will be objective and consistent in the scoring. Ideally, the potential for bias from the individual scoring the instrument will be completely eliminated by providing proper scoring instructions and other information necessary to guarantee an objective evaluation.

In some situations—for example, when self-scoring tests or group-based scoring mechanisms are used—participants are given the results immediately. In other situations, the results may not be known until later. In these situations, a method for providing the scores should be built into the evaluation plan unless it has been determined that participants will not know the scores. The worst

course of action is to promise that participants will receive their test scores and then deliver them late or not at all.

### Confronting Failure

Failure to pass the test may not be an issue, particularly if the data are collected informally through a self-assessment process. However, when more rigorous and formal methods are used and individuals do not demonstrate the required competency to pass the test, the issue of failures must be confronted. An important principle is to ensure that the test and the testing procedures are defensible. As described earlier, a test must be both reliable and valid, and the cut-off score for passing must be defensible. Written material should be developed to address these issues, and participants should be familiar with this material before they take the test. The outcomes and consequences of taking the test should also be discussed with the individuals. Repeating the test may be allowed, if appropriate, as long as all individuals receive the same treatment.

## Uses of Learning Data

Data must be used to add value and to improve processes. Several uses of learning data are appropriate; those described in this section are the most common.

*Ensuring that learning has been acquired.* Sometimes, knowing the extent and scope of learning is essential. Measuring learning, even informally, will provide input on this issue, indicating whether the learning component of the chain of input is successful.

*Providing individual feedback in order to build confidence.* Learning data, when provided directly to participants, support the learning process by providing reinforcement of correct answers and enhancing confidence.

*Improving a program.* Perhaps the most important use of learning data is improving a program. Learning data provide designers,

developers, facilitators, and team leaders with information and impetus for process improvement. Consistently low responses on certain learning measures may indicate that inadequate facilitation has been provided on that topic. Consistently low scores from all participants may indicate that the objectives and scope of coverage are misdirected or too ambitious.

*Evaluating facilitators*. In addition to reaction and planned action data, learning measures can be used to evaluate program leaders and facilitators, providing evidence of their success (or lack thereof). A facilitator has the responsibility of ensuring that participants learn the new skills and knowledge needed for program success. Learning measures reflect the degree to which the skills and knowledge have been acquired and internalized for application.

*Building a database*. In major programs that are repeated, building a database to track competency improvement, skills acquisition, or required knowledge may be helpful. Such data sets may be beneficial in indicating how one program compares with others. Over time, they can also be used to set expectations and judge success.

## Final Thoughts

This chapter discusses some of the key issues involved in measuring learning—an important ingredient in program success. Even if it is accomplished informally, learning must be assessed to determine the extent to which the participants in a program have learned new skills, techniques, processes, tools, and procedures. By measuring learning, facilitators and program leaders can ascertain the degree to which participants are capable of successfully implementing the program. Measuring learning provides an opportunity to quickly make adjustments and improvements in order to facilitate program success. While learning measures indicate potential success with implementation, measurement at the next level, application and implementation (Level 3), indicates actual progress.

## References

Lucia, A. D., and Lepsinger, R. *The Art and Science of Competency Models: Pinpointing Critical Success Factors in Organizations*. San Francisco: Pfeiffer, 1999.

Phillips, J. J., and Phillips, P. P. *Handbook of Training Evaluation and Measurement Methods*. (4th ed.) Woburn, Mass.: Butterworth-Heinemann, 2007.

Senge, P. *The Fifth Discipline: The Art and Practice of the Learning Organization*. New York: Random House, 1990.

Watkins, K. E., and Marsick, V. J. (eds.). *Creating the Learning Organization*. (J. J. Phillips, series ed.) Alexandria, Va.: ASTD, 1996.

# 7

# Measuring Application and Implementation

Many programs fail because of breakdowns in implementation. Participants just don't do what is expected, at the time expected, and as often as expected. Measuring application and implementation is critical to understanding the success of implementation. Without successful implementation, positive business impact will not occur and no positive return will be achieved.

This chapter explores the most common ways to evaluate the application and implementation of projects, processes, and programs. The possibilities vary from the use of questionnaires to observation to action planning. In addition to describing the techniques used to evaluate implementation, this chapter addresses the challenges and benefits of each technique.

## Why Measure Application and Implementation?

Measuring application and implementation is absolutely necessary for some programs. It provides the most critical data set because it helps evaluators understand the degree to which successful program implementation has occurred and the barriers and enablers that influence success.

### Obtain Essential Information

As Chapter One of *ROI Fundamentals* briefly discusses, the value of information to senior executives increases as evaluation progresses up the chain of impact from reaction (Level 1) to ROI (Level 5). Therefore, information concerning application and implementation (Level 3) is more valuable to clients than reaction and learning data (Levels 1 and 2). This fact does not discount the importance of evaluation at the first two levels but emphasizes the importance of moving up the chain of impact. Measuring the extent to which a program is implemented provides data about its success and about factors that can contribute to greater success as the program is fully implemented.

Level 1 and Level 2 measures occur during a program's early stages, when more attention is focused on participants' direct involvement in the program. Application and implementation measurement occurs after the program has been implemented. The data captured at this level reflect the program's success as participants apply the knowledge learned in the program to their job tasks. Essentially, this measure shows how well the system in which the participant works supports the transfer of knowledge to the job. Initial implementation of a program is the first step of a transition to a new state, behavior, or process. Understanding how successfully participants make the transition from applying skills in the program to applying them on the job requires measuring application and implementation.

### Track Program Focus

Because many projects and programs focus directly on implementation and application of new behaviors and processes, a program sponsor often speaks in these terms and has concerns about measures of implementation success. The sponsor of a major program designed to transform an organization will be concerned with implementation and application and thus will want to know the extent

to which key stakeholders adjust to and implement the desired new behaviors, processes, and procedures.

**Discover Problems and Opportunities**

If the chain of impact breaks at this level (application and implementation), little or no corresponding impact data will be available. Without impact, there is no return on the sponsor's investment. This breakdown most often occurs because participants in the program encounter barriers, inhibitors, or obstacles that deter implementation. A dilemma arises when reactions to the program are favorable and participants learn what is intended. Positive results at the reaction and learning levels set an expectation that knowledge transfer to the job will occur. Unfortunately, participants fail to overcome the barriers presented in the organization and don't accomplish what is necessary for success.

When a program goes astray, the first question usually asked is "What happened?" More important, when a program appears to add no value, the first question should be "What can we do to change its direction?" In responding to either question, it is important to identify the barriers to success, the problems with implementation, and the obstacles to application. At Level 3, implementation and application, these problems are addressed, identified, and examined. In many cases, the stakeholders directly involved in the process can provide important recommendations on making changes or using a different approach in the future.

When a program is successful, the obvious question is "How can we repeat this or improve it in the future?" The answer to this question is also found at Level 3. Identifying the factors that contribute directly to the success of a program is critical. Often, those same elements can be used to replicate the process and, possibly, to produce better results in the future. When key stakeholders identify critical elements and issues, they make the program successful and provide an important case history on what is necessary for success.

### Reward Effectiveness

Measuring application and implementation allows the sponsor and the program team to reward those who do the best job of applying the processes and implementing the program. Measures taken at this level provide clear evidence of success and achievement and also provide a basis for performance reviews. Rewards often have a reinforcing value, helping to keep employees on track and communicating strong encouragement for future improvement.

## Challenges

Collecting application and implementation data involves key challenges that must be addressed in order to attain a successful evaluation at this level.

### Linking Application with Learning

Application data should be linked closely with the learning data discussed in the preceding chapter. Program leaders need to know what has been accomplished, what has been done differently, and what activities have been implemented, all in comparison with what the participants learned to do in the program. This level of evaluation measures the extent to which participants accurately took what they learned and applied it to their jobs.

### Building Data Collection into the Program

Application data are collected after the program's implementation. Because of the time lag between program implementation and data collection, it is difficult to secure high-quality data and a large quantity of data. One of the most effective ways to ensure that the appropriate quality and quantity of data are collected is to build data collection into the program from the beginning. Data collection tools positioned as application tools must be built in as part of the implementation. By analogy, consider that many

software applications contain overlay software that shows a user performance profile. Essentially, the software invisibly tracks the user, capturing the steps, pace, and difficulties encountered while using the software. When the process is complete, a credible data set has been captured, simply because program leaders built it into the process at the beginning.

### Ensuring a Sufficient Amount of Data

Regardless of whether data were collected through questionnaires, action plans, interviews, or focus groups, poor response rates are a problem in most organizations. Motivating individuals to participate in the data collection process is a challenge. Ensuring that adequate amounts of high-quality data are available requires a serious effort to achieve adequate response rates.

When the program is intended to improve business impact and ultimately show a positive ROI, there is less emphasis on application and implementation measures; in many cases, they are omitted or slighted during the analysis. This is unfortunate as attention must be given to changing processes, procedures, and tasks, as well as to removing barriers in order to achieve the business goal. Doing things differently can result in substantial benefits, but knowing the degree to which things are done differently and the supporting or deterring influences is essential to guaranteeing those benefits in the future.

### Addressing Application Needs at the Outset

During needs assessment, one question that is asked is "What is being done (or not being done) on the job that's inhibiting [business measure]?" When this question is answered adequately, a connection is made between the solution and the business measure. When this issue is addressed, the activities or behaviors that need to change are identified, serving as the basis of the data collection. The bottom line is that too many evaluations focus either on impact measures, which define the business measure intended to be improved, or on learning measures, which uncover what people do

not know. More focus is needed at Level 3, which involves the tasks, processes, procedures, and behaviors that need to be in place for success on the job.

## Measurement Issues

In measuring program application and implementation, the key issues are largely similar to those encountered in measuring reaction and learning. There are slight differences due to the later time frame in which these data are collected.

### Methods

Methods for collecting data at Level 3 include traditional surveys and questionnaires, observation, interviews, and focus groups. Other powerful methods include action planning and follow-up sessions.

### Objectives

As at the other levels of evaluation, the starting point for Level 3 data collection is the objectives set for program application and implementation. Without clear objectives, collecting data would be difficult. Objectives define what activity is expected. They provide specific milestones that indicate when one part or all of the process has been implemented. Typical application objectives are as follows:

- At least 99.1 percent of software users will be following the correct sequences after three weeks of use.
- Within one year of program implementation, 10 percent of employees will submit documented suggestions for cutting costs.
- Ninety-five percent of high-potential employees will complete individual development plans within two years of program implementation.

- Forty percent of the city's homeless population will apply for special housing within one year of program launch.
- Eighty percent of employees will use one or more of the three cost containment features of the health care plan in the next round of annual enrollment.
- Fifty percent of conference attendees will follow up with at least one contact from the conference within six months.

### Areas of Coverage

To a certain extent, the areas of coverage for application and implementation measurement in a given program will parallel the areas that were identified for learning measurement. The later time frame for implementation data collection changes the measurement from a predictive measure to a post-program measure. The key point is that this level focuses on action, not on the ability to act (Level 2) and not on the consequences of acting (Level 4).

The sheer number of activities to measure can be mind-boggling. Action areas for application will vary from program to program, depending on the need for the program and the objectives and measures that represent program success. Table 7.1 shows some typical areas.

### Data Sources

The many sources of evaluation data were identified in Chapter Six of *ROI Fundamentals*. Essentially, all key stakeholders are potential data sources. Perhaps the most important sources of data are the users of the solutions—that is, those directly involved in the application and implementation of the project or program. Members of the program team or team leaders charged with the implementation may also be good sources of implementation data. In some cases,

**Table 7.1. Examples of Action Areas for Application of Program Learning**

| Action | Explanation | Example |
|---|---|---|
| Increase | Increasing an activity or action | Use negotiation skills with a certain level of frequency. |
| Decrease | Decreasing an activity or action | Decrease the number of times water temperature is checked during manufacturing. |
| Eliminate | Stopping a task or activity | Eliminate the formal follow-up meeting and replace it with a virtual meeting. |
| Maintain | Keeping the same level of activity for a particular process | Continue to monitor the manufacturing process on the same schedule previously used. |
| Create | Designing or implementing a new procedure, process, or activity | Create a procedure for resolving the differences between two divisions. |
| Use | Using a process, procedure, skill, or activity | Use communication skills in difficult customer service situations. |
| Perform | Performing a task, process, or procedure | Conduct a post-audit review at the end of each marketing cycle. |
| Participate | Becoming involved in an activity or program | Submit a suggestion for reducing R&D costs. |

**Table 7.1. Examples of Action Areas for Application of Program Learning (*Continued*)**

| Action | Explanation | Example |
|---|---|---|
| Enroll | Signing up for a process, program, or project | Enroll in a career advancement program. |
| Respond | Reacting to groups, individuals, or systems | Respond to customer inquiries within fifteen minutes. |
| Network | Facilitating relationships with others who are involved in or have been affected by a program | Continue networking with contacts on at least a quarterly basis. |

the best source of application data may be organizational records such as audit reports, time sheets, and databases.

### Timing

The timing of application data collection varies. Because application data collection is a follow-up activity after the program launch, the key issue is determining the best time for a post-implementation evaluation. The challenge is to analyze the nature and scope of the application and implementation in order to determine the earliest point by which a trend or pattern may evolve. At this point, the application of skills has become routine and the implementation is making progress. Determining this point is a judgment call. Collecting data as early as possible is important so that adjustments to the implementation, if necessary, can still be made. At the same time, evaluators must wait long enough so that behavior changes have a chance to occur and implementation makes enough progress to be observed and measured. In programs spanning a considerable length of time, measures may be taken at three- to six-month intervals. Using effective measures at well-timed intervals will provide

successive input on the progress of implementation and clearly show the extent of improvement.

Convenience and project constraints also influence the timing of data collection. If the participants are meeting to observe a milestone or a special event, the meeting provides an excellent opportunity to collect data. Sometimes, constraints are placed on data collection. Consider, for example, the time constraint that sponsors may impose. If they are anxious to have the data in order to make program decisions, they may request that data collection be moved to an earlier time than is ideal.

### Responsibilities

Measuring application and implementation usually requires many people to share the responsibility and the work. In Level 3 data collection, an important issue is determining who is responsible for following up on the progress of implementation. Application data collection may fall to personnel ranging from program staff and sponsors to external consultants. This matter should be addressed during the planning stages so that no misunderstanding arises about the distribution of responsibilities. More important, those who are responsible should fully understand the nature and scope of their role and what will be needed to collect the data.

## Data Collection Methods

Some techniques for collecting application and implementation data are easy to administer and provide quality data. Other techniques provide greater detail about program success but raise more challenges in administration.

### Questionnaires

Questionnaires have become a mainstream tool for collecting application and implementation data because of their flexibility, low cost, and ease of administration. The discussion of questionnaire

design in Chapter One can be applied to the development of questionnaires that measure application and implementation. One of the most difficult tasks is determining the issues to include in a follow-up questionnaire. Exhibit 7.1 presents an example of a questionnaire used to capture application, implementation, and impact data (Level 3 and Level 4 data).

A questionnaire like this might serve as the primary method of data collection for follow-up data. This example will be used to illustrate many of the issues involved in designing questionnaire items, especially items for collecting application (Level 3) and impact (Level 4) data.

*Progress with Objectives*

Sometimes, using a program's objectives to assess progress in the follow-up evaluation is helpful, as illustrated in question 1 in Exhibit 7.1. While program objectives are usually assessed during the program (when Level 1 data are collected), it can be beneficial to revisit the objectives after the participants have had an opportunity to apply what has been learned.

*Use of Program Materials and Handouts*

If participants have been provided with materials to use on the job, determining the extent to which these materials are used is important. This is particularly helpful when operating manuals, reference books, and job aids have been distributed and explained during the program and are expected to be used on the job. Question 2 in Exhibit 7.1 focuses on this issue.

*Application of Knowledge and Skills*

Knowing the level of improvement in the skills directly linked to the program is important (see question 3). A more detailed variation of this question lists each skill and asks participants to indicate their frequency of use and effectiveness of use for each. For many skills, frequent use soon after acquisition is critical so that the skills

## Exhibit 7.1. Sample Questionnaire

**Pharma Company**

*Note:* This example is used only to illustrate a sampling of questions that may be asked on a follow-up questionnaire. It is not intended to represent a document that is ready for implementation.

**Instructions**

Please complete this questionnaire as promptly as possible and return it to the address shown on the last page. To provide responses, you will need to reflect on the leadership development program and think about specific ways in which you have applied what you learned from each session. It may be helpful to review the materials from each session.

Please take your time as you provide responses. Accurate and complete responses are very important. You should be able to provide thorough responses in about 25 minutes.

Please be objective in providing responses. In no way will your name be linked to your input. Your questionnaire and action plan will be viewed only by a representative from an external firm, the ROI Institute. Specific responses or comments related to any individual will not be communicated to your employer.

Your responses will help determine the impact of this program and provide data for making adjustments. In exchange for your participation in this evaluation, a summary of the success of the entire class will be sent to you within two weeks. Three weeks later, you will receive a summary of the changes made to the program based on your input. Please make sure that your input is included along with that of your classmates.

Should you need clarification or more information, please contact your instructor, your representative, or a representative from the ROI Institute.

For each person returning the questionnaire, we are pleased to provide a copy of Marshall Goldsmith's book *What Got You Here Won't Get You There.* When you have returned the questionnaire, please let Crystal know, and she will be happy to send you a copy of the book. Thanks for your cooperation on this very important issue.

Jim Rogers
Chief Executive Officer

## Exhibit 7.1. Sample Questionnaire (*Continued*)

**Leadership Development Program Impact Questionnaire**

Are you currently in a supervisory or management role or capacity? Yes ❑ No ❑

1. Listed below are the objectives of the leadership program. After reflecting on the program, please indicate your degree of success in achieving these objectives. *Please check the appropriate response beside each item.*

| | No Success | Very Little Success | Limited Success | Generally Successful | Completely Successful |
|---|---|---|---|---|---|
| A. Apply the 11-step goal-setting process | ❑ | ❑ | ❑ | ❑ | ❑ |
| B. Apply the 12-step leadership planning process | ❑ | ❑ | ❑ | ❑ | ❑ |
| C. Identify the 12 core competencies of outstanding leaders | ❑ | ❑ | ❑ | ❑ | ❑ |
| D. Identify 10 ways to create higher levels of employee loyalty and satisfaction | ❑ | ❑ | ❑ | ❑ | ❑ |
| E. Apply the concept of deferred judgment in 5 scenarios | ❑ | ❑ | ❑ | ❑ | ❑ |
| F. Apply the creative problem-solving process to an identified problem | ❑ | ❑ | ❑ | ❑ | ❑ |
| G. Identify the 7 best ways to build positive relationships | ❑ | ❑ | ❑ | ❑ | ❑ |
| H. Given a work situation, apply the 4-step approach to deal with errors | ❑ | ❑ | ❑ | ❑ | ❑ |
| I. Practice 6 ways to improve communication effectiveness | ❑ | ❑ | ❑ | ❑ | ❑ |

2. Have you used the written materials since you participated in the program?
   Yes ❑ No ❑
   Please explain: _____

3. In the following result areas, please indicate the level of improvement during the last few months as influenced by your participation in the leadership program. *Check the appropriate response beside each item.*

## Exhibit 7.1. Sample Questionnaire (*Continued*)

| *Result Area* | No Opportunity to Apply | No Change | Some Change | Moderate Change | Significant Change | Very Significant Change |
|---|---|---|---|---|---|---|
| A. ORGANIZING | | | | | | |
| 1) Prioritizing daily activities | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| 2) Applying creative techniques | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| 3) Organizing daily activities | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| 4) Raising level of performance standards in area of responsibility | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| B. WORK CLIMATE | | | | | | |
| 1) Applying coaching | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| 2) Applying techniques or initiatives that influence motivational climate | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| 3) Implementing actions that influenced retaining people | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| 4) Implementing job enrichment opportunities for valued associates | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| 5) Implementing better control and monitoring systems | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| 6) Applying techniques that influenced better teamwork | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| C. PERSONAL OUTCOMES | | | | | | |
| 1) Realizing improved written communications | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| 2) Realizing improved oral communications | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| 3) Realizing greater self-confidence | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| 4) Working personal leadership plan | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |

## Exhibit 7.1. Sample Questionnaire (*Continued*)

4. List the three behaviors or skills from the above list that you have used most frequently as a result of the program.
   A) _____
   B) _____
   C) _____

5. What has changed about you or your work as a result of your participation in this program? (Specific behavior change such as increased delegation to employees, improved communication with employees, employee participation in decision making, improved problem solving, and so forth.)
   _____
   _____
   _____
   _____

6. How has Pharma Company benefited from your participation in the program? Please identify specific business accomplishments or improvements that you believe are linked to participation in this program. (Think about how the improvements actually resulted in influencing business measures, such as increased revenue, increased overall shipments, improved customer satisfaction, improved employee satisfaction, decreased costs, saved time, and so forth.)
   _____
   _____
   _____

7. Please define the measure(s) listed for Question 6. Be specific.
   _____
   _____

8. Indicate the amount of change in this measure since your participation in the program.
   _____
   Indicate the frequency of the measure.
   ❑ Daily ❑ Weekly ❑ Monthly ❑ Quarterly ❑ Annually

9. Think of specific ways to convert your accomplishments into a monetary value.
   Estimated unit monetary amount: $_____

10. What is the basis for the value listed in Question 9?_____
    ❑ Standard value
    ❑ Expert input (indicate who the expert is) _____
    ❑ Estimate (Indicate how you arrived at the value) _____

11. Calculate the total annual value of this measure, considering the change that has occurred following the program. Annualize that value, based on the frequency of the measure. Multiply the unit value by the amount of change times the needed adjustment for a whole year of change. _____

## Exhibit 7.1. Sample Questionnaire (*Continued*)

12. List factors other than the program that could have contributed to the improvement you listed in Question 7.

    _____

    _____

13. What percentage of the improvement above was actually influenced by the application of knowledge and skills from the leadership program?_____ %

14. What level of confidence do you place on the above estimations?
    _____% Confidence (0% = no confidence; 100% = certainty)

15. Do you think this leadership program represented a good investment for Pharma Company?
    Yes ❑ No ❑
    Please explain: _____

16. Indicate the extent to which you think your application of the knowledge, skills, and behavior learned from the leadership program had a positive influence on the following business measures in your own work or your work unit. *Please check the appropriate response beside each measure.*

| *Business Measure* | Not Applicable | Applies But No Influence | Some Influence | Moderate Influence | Significant Influence | Very Significant Influence |
|---|---|---|---|---|---|---|
| A. Work output | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| B. Quality | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| C. Cost control | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| D. Efficiency | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| E. Response time to customers | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| F. Cycle time of products | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| G. Sales | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| H. Employee turnover | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| I. Employee absenteeism | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| J. Employee satisfaction | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| K. Employee complaints | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| L. Customer satisfaction | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| M. Customer complaints | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |
| N. Other (please specify) | ❑ | ❑ | ❑ | ❑ | ❑ | ❑ |

## Exhibit 7.1. Sample Questionnaire (*Continued*)

17. Please cite specific examples or provide more details:

18. What additional benefits have been derived from this program?

19. What barriers, if any, have you encountered that have prevented you from using skills or behaviors gained in the leadership program? *Check all that apply.*
    - ❑ I have had no opportunity to use the skills.
    - ❑ I have not had enough time to apply the skills.
    - ❑ My work environment does not support the use of these skills or behaviors.
    - ❑ My supervisor does not support this type of program.
    - ❑ This material does not apply to my job situation.
    - ❑ Other (please specify):

    *If any of the above is checked, please explain, if possible.*

20. What enablers, if any, are present to help you use the skills or knowledge gained from this program? Please explain.

21. What additional support could be provided by management that would influence your ability to apply the skills and knowledge learned from the program?

22. What additional solutions do you recommend that would help to achieve the same business results that the leadership program has influenced?

**Exhibit 7.1. Sample Questionnaire (*Continued*)**

23. Would you recommend the leadership program to others?
    Yes ❑ No ❑
    Please explain. If no, why not? If yes, what groups or jobs, and why?

24. What specific suggestions do you have for improving this program?

25. Other comments:

become internalized. In our example, question 4 addresses the skill frequency issue in a more concise format.

*Changes in Work Activities*

Sometimes it is helpful to determine what specific work activities or processes have changed as a result of employees' participation in the program. Question 5 asks participants to explore how the application of skills learned in the program (listed earlier) has changed their work habits or processes.

*Improvements or Accomplishments*

Question 6 in Exhibit 7.1 begins a series of questions about business impact that are appropriate for most follow-up questionnaires. This question seeks specific accomplishments or improvements that are directly linked to the program, focusing on specific measurable successes that can be easily identified by the participants. Since this question is open-ended, providing examples that indicate the nature and range of responses requested may be helpful. However, when writing questions, consider that examples may also be constraining and may limit responses.

*Definition of the Measure*

Defining business measures precisely is important. Using general terms. such as quality or productivity or even sales. does not define a specific measure. A precise definition is necessary in order to provide an accurate accounting of improvement. Question 7 in Exhibit 7.1 allows participants to supply their definition of the business measure or measures they listed.

*Amount of Change*

In question 8, it is important for the participant to indicate how much the business measure listed earlier has changed. For example, if customer complaints are the measure, then the actual change in the number of complaints from before the program to after the program would be listed. In addition, the frequency of the measure is included. This information is important so that results can be annualized.

*Unit Value*

In order for a particular business measure to be used in ROI analysis, it must be assigned a unit monetary value. Many methods are available for assigning monetary value to data. Capturing this value directly from participants is important (see question 9).

Participants have three feasible sources for an estimate of unit value. First, a standard value for the measure may already be used within the organization. Fortunately, most measures that matter to an organization already have a unit value assigned to them. Second, an expert may be able to estimate a unit value. Finally, if the other two options are not available, the participant can estimate the unit value.

*Basis for Value*

To support the credibility of the estimated unit value, it is important to understand where it came from. Question 10 asks participants to indicate whether the unit value they are using is a standard value in the organization, whether it is based on expert input, or whether it is their own estimate. If the value comes from an expert, they are asked to indicate who the expert is. If it is an estimate, they are asked to show how the value was developed.

*Total Annual Impact*

Question 11 asks participants to show the total monetary value of the change that has occurred as a result of their participation in the program. The question asks participants to calculate the value of change for only one year. Participants can calculate this value by using the unit value, the amount of change, and the frequency of the data. For example, if the number of customer complaints has been reduced from 250 to 200 per month, and the unit value is $900 per complaint, then the total annual value is $50 \times \$900 \times 12$, which is $540,000.

*Other Factors*

When isolating the effects of a program on a business measure, it is important to think about other factors that might have caused the change. In question 12, participants are asked to list what other factors are in play that could have influenced the measure they listed.

*Improvements Linked with the Program*

The next question in the impact series (question 13) isolates the effects of the program. Participants estimate the percentage of the improvement that is directly related to the program. As an alternative, participants may be provided with various factors that may have influenced their results and asked to allocate percentages to each factor.

*Confidence Level*

To adjust for the uncertainty of the data provided in question 13, question 14 asks participants to offer a level of confidence for the estimate, expressed as a percentage from 0 to 100. This question allows participants to reflect their level of uncertainty about the estimating process.

*Perception of Investment in the Program*

Participants' views on the value of a program can be useful information. Question 15 of Exhibit 7.1 asks participants whether they perceive the program as an appropriate investment. Another option for a question on this topic is to present the actual cost of the program so that participants can respond from more accurate knowledge of the actual investment amount. It may be useful to express the cost as a per-participant cost. Also, the question can be divided into two parts—one reflecting on the investment of funds by the company and the other on the investment of the participant's time.

*Link with Output Measures*

Sometimes, it is helpful to determine the degree to which a program has influenced specific output measures, as shown in question 16. In some situations, a detailed analysis of the responses to such a question may reveal specifically which measures the program has influenced. However, when a program's effects are uncertain, it

may be helpful to list the business performance measures that may have been influenced by the program and seek input from the participants. The question should be worded so that the frame of reference is the time period after the program was conducted.

*Other Benefits*

In most programs, additional benefits—in particular, intangible benefits—will begin to emerge as the program is implemented. Participants should be asked to detail any benefits not represented elsewhere in the survey. In our example, question 18 is an open-ended query about additional benefits.

*Barriers*

A variety of barriers may detract from the successful application of the skills and knowledge learned in a program. Question 19 asks participants to identify such barriers. Some perceived barriers are listed, and the participants are asked to check all that apply. One alternative to the question in the exhibit is an open-ended question on this topic. Still another variation is listing the barriers with a range of responses that indicate the extent to which each barrier inhibited results.

*Enablers*

Just as important as barriers are enablers—issues, events, or situations that enable a process to be applied successfully on the job. Question 20 provides an open-ended question about enablers. The same options apply for this question as for question 19, only in reverse.

*Management Support*

For most programs, management support is critical to the successful application of newly acquired skills. At least one question on management support, such as question 21 in Exhibit 7.1, should

be included in the data collection instrument. Sometimes, this question is structured so that various descriptions of management support are detailed and participants are asked to check the one that applies to their situation. The information collected from questions about management support is very beneficial in helping to remove or minimize barriers.

*Other Solutions*

A program is only one of many potential solutions to a performance problem. If the needs assessment is faulty or if there are alternative approaches to developing the desired skills or knowledge, other potential solutions could be more effective in achieving the same success. In question 22, the participant is asked to identify other solutions that might be effective in obtaining the same or similar results.

*Target Audience Recommendations*

Sometimes, soliciting input about the most appropriate target audience for a program is useful. In question 23, participants are asked to indicate which groups of employees would most benefit from attending the program.

*Suggestions for Improvement*

For a wrap-up, participants are asked to provide suggestions for improving any part of the program. Illustrated in question 24, the open-ended structure is intended to solicit qualitative responses that can be used to make improvements.

---

The questionnaire in Exhibit 7.1 is quite comprehensive, soliciting a tremendous amount of data. All the data are needed to calculate the actual ROI and develop the full profile of data sets required for complete analysis. Obviously, many variations

on these questions could work well, and other questions could explore many other issues pertaining to program application and implementation.

### Interviews, Focus Groups, and Observation

Interviews and focus groups can be used during implementation or on a follow-up basis to collect data on implementation and application. Guidelines for designing and administering these instruments were covered in Chapter Three and will not be repeated here. Other resources also cover this area well (for example, Phillips and Phillips, 2007).

Observing participants on the job and recording changes in behavior and specific actions taken is another method of collecting Level 3 data. While observation is also used to collect learning data, a fundamental difference is that participants do not necessarily know they are being observed when observation is used to collect application data. Observation of participants is often used in sales and sales support programs. The observer may be a member of the program staff, the participant's manager, a peer, or an external resource such as a mystery shopper. Observation is most commonly (and, often, most practically) performed by a member of the program staff. Technology also provides tools to assist with observations. Audio recorders, video cameras, and computers play an important role in capturing application data.

### Action Plans

Action plans are a common follow-up approach. Participants are required to develop action plans as part of a program. The action plans contain the detailed steps necessary to accomplish specific objectives related to the program. The action planning process is one of the most effective ways to build participant support and the sense of ownership needed for successful program application and implementation. Level 3 data are collected during the follow-up

phase, when participants report on implementation of the action plans. Details of the design of action plans were discussed in Chapter Four, with an example provided in Exhibit 4.1.

## Barriers to Application

One of the important reasons for collecting application and implementation data is to uncover barriers and enablers. Although both groups are important, barriers can kill a program. Barriers must be identified, and actions must be taken to minimize, remove, or go around them.

Barriers are a serious problem in every program implementation. When they can be removed or minimized, the program can be implemented. When barriers are identified, they become important reference points for change and improvement. Typical barriers that can stifle the success of programs include the following:

- My immediate manager does not support the program.
- The culture in our work group does not support the program.
- We have no opportunity to use the program skills, knowledge, or information.
- We have no time to implement the program.
- Technology was not available for the program.
- Resources are not available to implement the program.
- We didn't see a need to implement the program.
- Another program got in the way.
- The program is not appropriate for our work unit.
- My job changed, and this no longer applies.

The important point is to identify all barriers and to use the data in meaningful ways to make them less of a problem.

## Uses of Application Data

Data become meaningless if they are not used properly. As evaluation moves up the chain of impact, the data become more valuable in the minds of sponsors, key executives, and others who have a strong interest in a program. Although data can be used in dozens of ways, the principal uses for Level 3 data after they are collected are as follows:

- To report results and review them with various stakeholders
- To adjust program design and implementation
- To identify and remove barriers
- To identify and enhance enablers
- To recognize individuals who have contributed to program success
- To reinforce in current and future program participants the value of desired actions
- To improve management support for programs
- To market future programs

## Final Thoughts

Measuring application and implementation is critical in determining whether a program has been successful. This essential type of measurement not only determines the success achieved but also identifies areas in which improvement is needed and in which

success can be replicated in the future. Understanding success in application is important in providing evidence that business needs have been met, but it is only through measurement at Level 4, impact and consequences, that a direct link between a program and business impact can be made.

## Reference

Phillips, J. J., and Phillips, P. P. *Handbook of Training Evaluation and Measurement Methods*. (4th ed.) Woburn, Mass.: Butterworth-Heinemann, 2007.

# 8

# Measuring Impact and Consequences

Most clients regard business impact data as the most important data type because of its connection to business success. In many programs, inadequate performance in one or more business measures (the business need) is what led executives to initiate the program. Business impact evaluation data close the loop by showing a program's success in meeting business needs. This chapter examines a variety of business impact measures and the specific processes needed to collect the measurement data within a program. First, however, this chapter addresses the reasons that impact data are measured.

## Why Measure Business Impact?

Several rationales support the collection of business impact data related to a program.

### Impact Data Provide Higher-Level Information on Performance

If we go with the assumption that higher-level data create more value for key stakeholders, business impact measures offer a more valuable data set than measures at Level 1, 2, or 3. Why are Level 4 data more valuable? Impact data are the consequence of application and implementation of a program. They represent

the bottom-line measures that are positively influenced when a program is successful.

The chain of impact can be broken at Level 4, and unfortunately, in many programs, it is. If the program does not drive business impact data, results at lower levels may be less than satisfactory. However, in some cases, the program is successful at the lower levels but fails at Level 4. Participants may react positively to the program, may learn successfully to implement the program, and may follow the correct implementation steps or use the skills needed to implement the program. However, if the business impact measure that was thought to be influenced by the program does not change, the program does not add value. What could cause this? There are two possibilities. First, the program may not have been properly aligned with the business, which would mean that the program is not the right solution. Although the program may have been well implemented, it has driven activity and not results. The second possibility is that factors other than those addressed by the program are driving the business measure. Although the program may be connected to the business measure, other influences may be negatively affecting the same measure. Thus, at first glance, it may appear that the program has no value, but in reality it may have. This brings into focus the importance of isolating the effects of a program. The business data may be disappointing, but it is possible that they would be even more disappointing without the program. The important process of isolating the effects of a program is presented in book three of this series, *Isolation of Results*.

**Impact Data Represent the Business Driver of a Program**

For most programs, business impact data represent the initial drivers of the program. The problem of deteriorating or poorer-than-expected performance or the opportunity for improvement of performance on a business measure usually leads to initiation of a program. If the business needs defined by a business measure are driving a program, then the key measure for evaluating the

program is that business measure. The extent to which that measure changes is the principal determinant of success.

### Impact Data Provide Value for Sponsors

From the perspective of a sponsor, business impact data represent key measures of a program's payoff. These key measures are the ones often desired by the sponsor and the ones that the sponsor wants to see changed or improved. They often represent hard, indisputable facts about business performance, which is critical at the business unit or operating unit level of the organization. Business impact leads to "the money"—the actual return on investment in the program. Without credible business impact data linked directly to the program, it would be difficult, if not impossible, to establish a credible monetary value for the program. These business realities make this level of data collection one of the most critical.

### Impact Data Are Easy to Measure

One unique feature of business impact data is that they are often easy to measure. Hard and soft data measures at this level often correspond to key measures that are plentiful throughout an organization. It is not unusual for an organization to have hundreds or even thousands of measures reflecting specific business impact items. The challenge is to connect the objectives of the program with the appropriate business measures. This linkage begins during the needs assessment process, continues as program objectives are developed, and is made clear during the evaluation process with the step to isolate the program's impact on Level 4 measures.

## Effective Impact Measures

Any process, item, or perception can be measured, and such measurement is critical to Level 4 analysis. If the program focuses on solving a problem, preventing a problem, or seizing an opportunity, the measures are usually identifiable. The important point is that

the measures are present in the system, ready to be captured for this level of analysis. The challenge is to define the measures and to find them economically and swiftly.

### Hard Data Measures

In choosing desired measures, a distinction between hard data and soft data is often used to categorize the various types of business measures. Hard data are measures of improvement presented in the form of quantitative, undisputed facts that are usually gathered within functional areas throughout an organization. These are the most desirable type of data because they are easy to quantify and are easily converted to monetary values. The fundamental criteria for gauging the effectiveness of an organization are hard data items such as revenue, productivity, and profitability, as well as measures of quality, costs, and time.

Hard data are objective and credible measures of an organization's performance. Hard data are usually grouped in four categories, as shown in Table 8.1. These categories—output, quality, costs, and time—are typical performance measures in any organization.

Hard data from a particular program often involve improvements in the output of a work unit, section, department, division, or entire organization. Every organization, regardless of type, must have basic measures of output, such as number of patients treated, students graduated, tons produced, or packages shipped. Since these values are monitored, changes can easily be measured by comparing outputs before and after a program.

Quality is a very important category of hard data. If quality is a major priority for an organization, processes are likely in place to measure and monitor quality. The rising prominence of quality improvement processes (such as total quality management, continuous quality improvement, and Six Sigma) has contributed to tremendous recent successes in pinpointing an organization's proper quality measures, as well as assigning monetary values to them.

Table 8.1. Examples of Hard Data

| Output | Quality | Costs | Time |
|---|---|---|---|
| Units produced | Failure rates | Shelter costs | Cycle time |
| Tons manufactured | Dropout rates | Treatment costs | Equipment downtime |
| Items assembled | Scrap | Budget variances | Overtime |
| Money collected | Waste | Unit costs | On-time shipments |
| Items sold | Rejects | Cost by account | Time to program completion |
| New accounts generated | Error rates | Variable costs | Processing time |
| Forms processed | Rework | Fixed costs | Supervisory time |
| Loans approved | Shortages | Overhead costs | Time to proficiency |
| Inventory turnover | Product defects | Operating costs | Learning time |
| Patients visited | Deviation from standard | Program cost savings | Adherence to schedules |
| Applications processed | Product failures | Accident costs | Repair time |
| Students graduated | Inventory adjustments | Program costs | Efficiency |
| Tasks completed | Time card corrections | Sales expense | Work stoppages |
| Output per hour | Incidents | | Order response time |
| Productivity | Compliance discrepancies | | Late reporting |
| Work backlog | Agency fines | | Lost-time days |
| Incentive bonus | | | |
| Shipments | | | |
| Completion rate | | | |

Cost is another important category of hard data. Many projects and programs are designed to lower, control, or eliminate the cost of a specific process or activity. Achieving cost targets has an immediate effect on the bottom line. Some organizations focus narrowly on cost reduction. For example, consider Wal-Mart, whose tagline is "Always low prices. Always." All levels of the organization are dedicated to lowering costs on processes and products and passing the savings along to customers.

Time is a critical measure in any organization. Some organizations gauge their performance almost exclusively in relation to time. When asked what business FedEx is in, company executives say, "We engineer time."

### Soft Data Measures

Soft data are probably the most familiar measures of an organization's effectiveness, yet collecting them can present a challenge. Values representing attitude, motivation, and satisfaction are examples of soft data. Soft data are more difficult to gather and analyze than hard data; therefore, they are used when hard data are not available or to supplement hard data. Soft data are also more difficult to convert to monetary values, a process that requires subjective methods. Performance measurements using soft data are less objective than those using hard data and are usually behavior-related, yet organizations place great emphasis on them.

Improvements in soft data measures represent important business needs, but many organizations omit them from the ROI equation due to the challenges in converting them to money. However, soft data measures can contribute to economic value to the same extent as hard data measures. Table 8.2 shows common examples of soft data by category. The key is not to focus too much on the hard versus soft data distinction. A better approach is to consider data as tangible or intangible.

**Table 8.2. Examples of Soft Data**

| | |
|---|---|
| **Work Habits** | **Customer Service** |
| Excessive breaks | Customer complaints |
| Tardiness | Customer satisfaction |
| Visits to the dispensary | Customer dissatisfaction |
| Violations of safety rules | Customer impressions |
| Communication breakdowns | Customer loyalty |
| | Customer retention |
| **Work Climate and Job Satisfaction** | Lost customers |
| Grievances | **Employee Development and Advancement** |
| Discrimination charges | Promotions |
| Employee complaints | Capability |
| Job satisfaction | Intellectual capital |
| Organization commitment | Requests for transfer |
| Employee engagement | Performance appraisal ratings |
| Employee loyalty | Readiness |
| Intent to leave the organization | Networking |
| Stress | |
| **Initiative and Innovation** | **Image** |
| Creativity | Brand awareness |
| Innovation | Reputation |
| New ideas | Leadership |
| Suggestions | Social responsibility |
| New products and services | Environmental friendliness |
| Trademarks | Social consciousness |
| Copyrights and patents | Diversity |
| Process improvements | External awards |
| Partnerships and alliances | |

### Tangible Versus Intangible Measures

The fundamental difference between hard and soft data involves the objectivity in their measurement. The key is to remember that, ultimately, all roads lead to hard data. Although creativity

may be categorized as a form of soft data, a creative workplace can develop new products or new patents, which leads to greater revenue—which clearly is a hard data measure. Although it is possible to convert soft data to monetary value, it is often more realistic and practical to leave them in nonmonetary form. This decision is based on considerations of credibility and the cost of the conversion.

According to the Guiding Principles of the ROI Methodology, an intangible measure is defined as a measure that is intentionally not converted to a monetary value. If a soft data measure can be credibly converted to a monetary amount, using minimal resources, it is considered tangible, reported as a monetary value, and incorporated in the ROI calculation. If a data item cannot be credibly converted to money, using minimal resources, it is listed as an intangible measure. Therefore, the key difference between measures in terms of ROI calculation is not whether they represent hard or soft data but whether they are tangible or intangible. In either case, they are important contributions toward measurement of the desired payoff and of important business impact data.

### Impact Objectives

Impact objectives indicate key business measures that should improve as the application and implementation objectives are achieved. Following are some typical impact objectives:

- Grievances should be reduced from three per month to no more than two per month at the Golden Eagle tire plant within one year of program launch.
- Tardiness at the Newbury foundry should decrease by 20 percent within the next calendar year.
- The average number of product defects should decrease from 214 to 153 per month at all Amalgamated Rubber

extruding plants in the Midwest region by the end of the current fiscal year.

- The company-wide job satisfaction index should rise by 2 percent during the next calendar year.
- There should be a 10 percent increase in Pharmaceuticals Inc. brand awareness among physicians during the next two years.
- The dropout rate for high school students in the Barett County system should decrease by 5 percent within three years.

Impact objectives are critical to measuring business performance because they define the ultimate expected outcome of the project. They describe the business unit performance that should result from the project. Above all, impact objectives emphasize achievement of the bottom-line results that key client groups expect and demand.

**Linking Specific Measures to Programs**

An important issue that often surfaces when considering ROI applications is how to understand which specific measures are driven by specific programs. Although no standard answers are available, Table 8.3 summarizes some typical payoff measures for specific types of programs. The measures are quite broad for some programs. For example, a reward systems program can pay off in a variety of measures, such as improved productivity, enhanced sales and revenues, improved quality, cycle-time reduction, and even direct cost savings. Essentially, the reward systems program should drive the measure that the rewards are designed to influence. In other programs, the influenced measures are quite narrow. For example, in labor-management cooperation programs, the payoffs are typically in reduced grievances, fewer work stoppages, lower absenteeism, and improved employee satisfaction. Orientation programs

**Table 8.3. Typical Impact Measures in ROI Applications**

| Program | Key Impact Measurements |
|---|---|
| Absenteeism control or reduction | Absenteeism, customer satisfaction, job satisfaction, stress |
| Business coaching | Productivity, output, quality, time savings, efficiency, costs, employee satisfaction, customer satisfaction |
| Career development and career management | Turnover, promotions, recruiting expenses, job satisfaction |
| Communications | Errors, stress, conflicts, productivity, job satisfaction |
| Compensation plans | Costs, productivity, quality, job satisfaction |
| Compliance | Penalties and fines, charges, settlements, losses |
| Diversity | Turnover, absenteeism, complaints, allegations, legal settlements, losses |
| E-Learning | Cost savings, productivity improvement, quality improvement, cycle times, error reductions, job satisfaction |
| Employee benefits plans | Costs, time savings, job satisfaction |
| Employee relations | Turnover, absenteeism, job satisfaction, engagement |
| Gainsharing plans | Production costs, productivity, turnover |
| Labor-management cooperation | Work stoppages, employee grievances, absenteeism, job satisfaction |
| Leadership development | Productivity, output, quality, efficiency, cost savings, time savings, employee satisfaction, engagement |
| Marketing and advertising | Sales, market share, customer loyalty, cost of sales, wallet share, customer satisfaction, brand recognition |
| Meeting planning | Sales, productivity, output, quality, time savings, job satisfaction, customer satisfaction |
| Orientation and on-boarding | New hire turnover, training time, productivity |

**Table 8.3. Typical Impact Measures in ROI Applications (*Continued*)**

| Program | Key Impact Measurements |
|---|---|
| Personal productivity and time management | Time savings, productivity, stress reduction, job satisfaction |
| Procurement | Costs, time savings, quality, stability, schedule |
| Project management | Time savings, quality improvement, budgets |
| Public policy | Time savings, cost savings, quality, stakeholder satisfaction, image |
| Public relations | Image, branding, customer satisfaction, investor satisfaction |
| Recruiting sources | Costs, yield, early turnover |
| Retention management | Turnover, engagement, job satisfaction |
| Safety incentive plan | Accident frequency, accident severity, first aid treatments |
| Selection process | New hire turnover, training time, productivity |
| Self-directed teams | Productivity, output, quality, customer satisfaction, turnover, absenteeism, job satisfaction |
| Sexual harassment prevention | Complaints, turnover, absenteeism, employee satisfaction |
| Six Sigma | Defects, rework, response times, cycle times, costs |
| Skill-based pay | Labor costs, turnover, absenteeism |
| Strategy and policy | Productivity, output, sales, market share, customer service, quality or service levels, cycle times, cost savings, job satisfaction |
| Stress management | Medical costs, turnover, absenteeism, job satisfaction |
| Technical training (job-related) | Productivity, sales, quality, time, costs, customer service, turnover, absenteeism, job satisfaction |
| Technology implementation | Cycle times, error rates, productivity, efficiency, customer satisfaction, job satisfaction |
| Wellness and fitness | Turnover, medical costs, accidents, absenteeism |

typically pay off in measures of early turnover (turnover in the first ninety days of employment), initial job performance, and productivity. The measures that are influenced depend on the objectives and the design of the program.

Table 8.3 also illustrates the immense number of applications of the ROI Methodology and the even larger set of measures that can be driven or influenced. In most of these situations, assigning monetary values to the impact measures (so that the benefits of a given program can be compared with the costs) and developing the ROI become reasonable tasks.

A word of caution: Presenting specific measures linked to a typical program may give the impression that these are the only measures influenced. In practice, a given program can have many outcomes, and this can make calculation of the ROI a difficult process. The good news is that most programs are driving business measures. The monetary values are based on what is being changed in various business units, divisions, regions, and individual workplaces. These are the measures that matter to senior executives. The difficulty often comes in establishing that a connection to the program exists. This connection is deduced through a variety of techniques that isolate the effects of the program on particular business measures, as will be discussed in *Isolation of Results*, the third book in this series.

**Sources of Impact Data**

The potential sources of impact data are diverse. Many data items come from routine reporting systems in the organization. In many situations, poor performance in regard to these items has led to the perceived need for the program. A vast array of documents, systems, databases, and reports can be consulted in selecting the specific measure or measures to be monitored throughout a program. Impact data sources include quality reports, service records, suggestion systems, and employee engagement data.

Some evaluators assume that corporate data sources are scarce because the data are not readily available to them. However, data

can usually be located by investing a small amount of time. Rarely do new data collection systems or processes need to be developed in order to gather data that represent the business needs of an organization.

In searching for the proper measures to connect to the program and to identify business needs, it is helpful to consider all possible measures that could be influenced. Sometimes, collateral measures move in harmony with the program. For example, efforts to improve safety may also improve productivity and increase job satisfaction. Weighing adverse impacts on certain measures may also help. For example, when cycle times are reduced, quality may suffer; or when sales increase, customer satisfaction may deteriorate. Finally, evaluators must anticipate unintended consequences and capture them as other data items that might be connected to or influenced by the program.

## Data Collection Methods

As with the previous levels of measurement, there are a variety of ways by which we can collect Level 4 business impact data. Many of these techniques are also used to collect data for lower levels of evaluation and are described in detail in Chapter Four. This section serves as a review of those techniques and describes how they are used with Level 4 measurement.

### Monitoring Business Performance Data

Data are available in every organization to measure business performance. Monitoring performance data enables management to measure performance in terms of output, quality, costs, time, job satisfaction, customer satisfaction, and other measures. In selecting sources of data for evaluation, the first consideration should be existing databases, reports, and scorecards. In most organizations, performance data will be available that are suitable for measuring improvement resulting from a program. If such data are not

available, additional record-keeping systems may have to be developed for measurement and analysis. At this point, the question of economics surfaces. Is it economical to develop the record-keeping systems necessary to evaluate a program or project? If the costs will be greater than the expected benefits, developing those systems is pointless.

*Identify Appropriate Measures*

Existing performance measures should be thoroughly researched to identify those most closely related to the proposed objectives of the program. Often, several performance measures are related to the same item. For example, the efficiency of a production unit can be measured in several ways:

- Number of units produced per hour
- Number of units produced on schedule
- Percentage of equipment used
- Percentage of equipment downtime
- Labor cost per unit of production
- Overtime required per unit of production
- Total unit cost

Each of these measures the effectiveness or efficiency of the production unit in its own way. All related measures should be reviewed to determine those most relevant to the program.

*Convert Current Measures to Usable Ones*

As discussed in Chapter Four, it is sometimes necessary to convert existing measures to usable measures. To do this, performance measures are integrated with other data. An example of this process is described in Chapter Four.

*Develop New Measures*

In some cases, data needed to measure the effectiveness of a program are not available, and new data are needed. This need can be met by working with the client or others in the organization to develop record-keeping systems. The first step is to identify the department that will develop the measurement system. This department is typically the department that "owns" the measure. For example, if absenteeism is a problem, which job category represents the position for which absenteeism is the problem? Determining the job category will often assist in isolating the department or function that should track the measure. If the problem is cross-organizational (meaning that absent employees are an issue across multiple job categories) and it is determined that the measure should be monitored from a corporate perspective, then it may be prudent to house the measurement system in the human resources or accounting and finance department.

Next, the person(s) who will actually monitor the measure should be identified. Third, it must be decided where the measures will be recorded—a database, a paper-based system, or some other system. Will there be a journal or a record book, or will automation be used to track the measure? Finally, what forms will be used to support various departments in tracking absenteeism in their function?

## Action Plans

In Chapter Seven, action plans were presented as a tool to capture application and implementation data. Action plans are also a useful tool for capturing business impact data. The business impact data gathered through use of an action plan are more focused and credible than those gathered through a questionnaire. The basic design principles and issues involved in developing and administering action plans to gather business impact data are the same as those discussed in Chapter Four. However, a few issues unique to

collecting business impact and ROI data are presented here. The following steps are recommended when an action plan is developed and implemented to capture business impact data and to convert the data to monetary values.

*Set Goals and Targets*

An action plan can be developed to focus directly on business impact data. Participants develop an overall objective for the plan, which is usually the primary objective of the program. In some cases, a program may have more than one objective, which requires additional action plans—one for each objective. In addition to the objective, the improvement measure and current levels of performance are identified. Providing this information on objectives requires the participant to anticipate the application and implementation of the program and to set goals for specific performances that can be realized.

The action plan is completed during program implementation, often with the input, assistance, and facilitation of the program team. The evaluator or program leader approves the plan, indicating that it meets the requirements of being specific, motivational, achievable, realistic, and time-based (SMART). The plan can be developed in one to two hours and often begins with action steps related to implementation of the program. These action steps are Level 3 activities that detail the application and implementation of learning from the program. All these steps build support for and are linked to business impact measures.

*Define the Unit of Measure*

The next important issue is defining the unit of measure for the desired change in performance. In some cases, more than one measure may be used, resulting in additional action plans—again, one for each measure. The unit of measure is necessary in order to break the process into simple steps so that its ultimate value can be determined. The unit may be a unit of output data—such as one unit manufactured or one package delivered—or it may be a unit of

sales and marketing data—such as $1 of sales revenue or a 1 percent increase in market share. If quality is being measured, the unit may be one reject, one error, or one defect. Time-based units are usually measured in minutes, hours, days, or weeks. Other units are specific to their particular type of data, such as one grievance, one complaint, one absence, or one person receiving welfare payments. The important point is to break impact data into the simplest terms possible.

*Place a Monetary Value on Each Improvement*

During program implementation, participants are asked to locate, calculate, or estimate the monetary value of each improvement outlined in their plan. The unit monetary value is determined through a variety of methods, including standard values, expert input, external databases, and estimates. The process to be used in arriving at the unit monetary value is described in the instructions of the action plan. When the actual improvement occurs, participants use the unit monetary values to calculate the annual monetary benefits of the plan. To be effective in this step of assigning monetary values to improvements, it is helpful to understand the ways in which values can be assigned to the data, which are discussed in *Data Conversion,* the fourth book of this series.

In the worst-case scenario, participants are asked to estimate unit monetary values themselves, although use of standard values and consultation with an expert are better courses of action. When it is necessary for participants themselves to make the calculations, they must explain the assumptions they made in their calculations.

*Implement the Action Plan*

Participants implement their action plan during program implementation, which often lasts for weeks or months following the launch of a program. Participants complete the activities described in the action plan, and business impact results are achieved.

*Document Specific Improvements*

At the end of the specified follow-up period—usually three months, six months, nine months, or one year—participants indicate the specific improvements that have been achieved, usually expressed as a daily, weekly, or monthly amount that represents the actual amount of change observed, measured, or recorded. Participants must understand the need for accuracy as data are recorded. In most cases, only the changes are recorded because these are the amounts needed to calculate the monetary value of the program. In other cases, data from before and after program implementation may be recorded, allowing the evaluator to calculate the difference.

*Isolate the Effects of the Program*

As discussed in Chapter Four, while the action plan may have been initiated because of the program, the improvements reported on the action plan may have been caused or influenced by factors other than those put in play by the program. Consequently, the program should not be given full credit for the improvement. For example, an action plan to implement a new computer system in a division may be given only partial credit for a business improvement because other variables may have affected the impact measures. Although there are several ways to isolate the effects of a program, participant estimation is usually most appropriate in the action planning process, because participants will typically be focusing on measures specific to their work rather than targeting improvement in one specific, universal measure. Consequently, participants are asked to estimate the percentage of the improvement that is related to a particular program. This information can be requested on the action plan form or in a follow-up questionnaire. Sometimes it is beneficial to precede the request for an estimate with a request that the participant identify the entire range of factors that could have influenced the results. This encourages participants to think through the relationships among all the factors before allocating a portion of the results to the program under evaluation.

*Provide a Confidence Level for Estimates*

The process of isolating the amount of improvement related to the program is not always precise. Participants are asked to indicate their level of confidence in their estimate. By using a scale of 0 to 100 percent in which 0 indicates that the values are completely false and 100 percent indicates that the values are absolutely certain, participants can express their uncertainty about their estimate.

*Collect Action Plans at Specified Time Intervals*

Because a high response rate is essential to a credible evaluation process, several steps may be necessary to ensure that action plans are completed and returned. Participants usually see the importance of the process and develop their plans in detail at the beginning of the program. Some organizations send follow-up reminders by mail or e-mail; others phone participants to check on their progress. Still others offer assistance in developing the final plan. These steps may require expending additional resources, the value of which must be weighed against the importance of having more data. Specific ways to improve response rates for questionnaires were discussed in Chapter One, but many of the techniques apply to action plans, too.

*Summarize the Data and Calculate the ROI*

If developed properly, each action plan will include annualized monetary values associated with improvements. Also, each individual should have indicated the percentage of the improvement that is directly related to the program. Finally, each participant should have provided a confidence percentage to reflect his or her uncertainty about the process and the subjective nature of some of the data provided.

---

Because this process involves estimates, it may appear to be inaccurate, although adjustments during analysis can make the process more credible and more accurate. These adjustments reflect

the Guiding Principles that form the basis of the ROI Methodology. Exhibit 8.1 shows an example of a completed action plan with the impact data reported.

### Performance Contracts

Another technique for collecting business impact data is the performance contract. This technique is described in Chapter Four. As described, a performance contract is a slight variation on the action plan. Based on the principle of mutual goal setting, a performance contract is a written agreement between a participant and the participant's manager. The participant agrees to improve performance in an area of mutual concern related to the program. The agreement establishes a goal for the employee to accomplish during the program or after the program's completion. It will sometimes include the manager's commitment to the employee to ensure success. The agreement details what is to be accomplished, at what time, and with what results. Exhibit 8.2 presents a simple performance contract.

### Questionnaires

As described in previous chapters, the questionnaire is one of the most versatile data collection tools and can be used to collect Level 1, 2, 3, and 4 data. Essentially, the design principles and content issues are the same as at other levels, except that questionnaires developed for a business impact evaluation will include additional questions to capture data specific to business impact. A detailed example of such a questionnaire was included in the preceding chapter.

The use of questionnaires for impact data collection brings both good news and bad news. The good news is that questionnaires are easy to implement and low in cost. Data analysis is efficient, and the time required for participants to provide the data is often minimal, making questionnaires among the least disruptive of data collection methods. The bad news is that the data can be distorted

Exhibit 8.1. Example of an Action Plan

## Action Plan

Name: *John Mathews* Instructor's Signature: _____ Follow-Up Date: *1 September*

Objective: *Reduce team's weekly rate of absenteeism* Evaluation Period: *March to September*

Improvement Measure: *Rate of absenteeism* Current Performance: *8%* Target Performance: *5%*

### Action Steps

| | |
|---|---|
| 1. *Meet with team to discuss reasons for absenteeism, using problem-solving skills.* | *10 March* |
| 2. *Review absenteeism records for each employee; look for trends and patterns.* | *20 March* |
| 3. *Counsel problem employees to help them correct habits and explore opportunities for improvement.* | |
| 4. *Conduct a brief performance discussion with an employee returning to work after an unplanned absence.* | |
| 5. *Provide recognition to employees who have perfect attendance.* | *31 March* |
| 6. *Follow up on each discussion; discuss improvement or lack of improvement, and plan other action.* | |
| 7. *Monitor improvement and provide recognition when appropriate.* | |

*Intangible Benefits*: *Less stress, greater job satisfaction*

### Analysis

A. What is the unit of measure? *One absence*

B. What is the value (cost) of one unit? *$41.00*

C. How did you arrive at this value? *Standard value*

D. How much did the measure change during the evaluation period? (monthly value) *2.5%*

E. List the other factors that could have caused this. *Job market and new discipline policy.*

F. What percentage of this change was actually caused by this program? *65%*

G. What level of confidence do you place on the above information? (100% = certainty; 0% = no confidence) *80%*

Comments: *Great program—it kept me on track with this problem.*

## Exhibit 8.2. Example of Simple Performance Contract

**ROI RESOURCE CENTER™**

**Performance Contract**

To: Carrie Luke

From: Patti Phillips

Date: April 5

We are happy to have your talents added to the staff. Your expertise will undoubtedly promote the program's overall success that is important to ROI Resource Center. To ensure that we are all in agreement about deliverables and timelines, this document summarizes the expected performance from the upcoming program.

The program we plan to implement was designed specifically to help us develop innovative products that support clients' evaluation practices.

**Desired performance:** 10 new products by year end

**Current performance:** 5 new products annually

The program will be considered successful to the extent we see improvement in the performance from the current to desired level.

In order to complete your part of the project, you will receive the following from us by the dates indicated:

| Items (that your organization will provide) | Date |
|---|---|
| A copy of the entire performance improvement action plan. | April 15 |
| Draft product criteria | April 30 |
| Focus group participant list | May 15 |
| Final product selection matrix | May 30 |
| Product selection | June 15 |
| Focus group logistics | June 30 |

We expect that you will deliver the following items on the date(s) indicated:

| Deliverables (items that your organization will need) | Date |
|---|---|
| Comments on product criteria | May 10 |
| Suggestions and quotes for new products | June 15 |
| Sample products to demo in focus group | July 10 |
| Conduct focus group | July 15 |
| Focus group results | July 20 |

We agree to provide the deliverables specified above on the dates noted.

Signature _____ Director_____

Date _____ Date_____

and inaccurate, and data are sometimes missing. Poor design, lack of understanding of program objectives, and misalignment between questions and measures can undermine attempts to gather good data. The challenge is to take all the steps necessary to ensure that questionnaires are complete, accurate, and clear, and that they are returned with all the necessary data.

Unfortunately, questionnaires are the weakest method of data collection due to the subjectivity inherent in the use of this instrument. Paradoxically, they are the most commonly used because of their advantages. Of the first one hundred case studies published on the ROI Methodology, roughly half used questionnaires as a method of data collection. Questionnaires are popular, convenient, and low in cost, and they have become a way of life. The challenge is to improve them. The philosophy of the ROI Methodology is to take processes that represent the weakest method and make them as credible as possible. Here the challenge is to make questionnaires credible and useful by ensuring that they request all the data needed, that participants provide accurate and complete data, and that return rates are in at least the 70 to 80 percent range.

The reason that return rates must be high is explained in Guiding Principle 6 of the ROI Methodology: no data, no improvement. If an individual provides no improvement data, it is assumed that the person had no improvement to report. This is a very conservative principle, but it is necessary to bring credibility to the evaluation process. Consequently, using questionnaires will require effort, discipline, and personal attention in order to ensure proper response rates. Chapter One presented many ways to ensure high response rates for Level 1 data collection. The same techniques should be considered for Level 4 data collection. It is helpful to remember that questionnaires are the least preferred method for collecting Level 4 data and should be used only when other methods do not work (that is, when business performance data cannot be easily monitored, when action plans are not feasible, or when performance contracting is not suitable).

## Final Thoughts

This chapter has explored the issues of collecting impact data, illustrating the common ways in which this important data set is captured. The good news is that these data are very common and readily available in virtually every organization. They can usually be collected by simply monitoring the organization's record-keeping systems. If not, action plans, performance contracts, and questionnaires can be used to capture this data set.

# 9

# Selecting the Proper Data Collection Method

Confused? Maybe so. There are many ways to collect data, leaving some people in a state of confusion as to which data collection method is most appropriate for their situation. This brief chapter will provide some practice in the use of the data collection tools described in this book. It also provides guidelines to consider when selecting the best technique(s) for collecting data when evaluating your programs.

## Matching Exercise

Perhaps an exercise will illustrate some of the key differences among the data collection methods. While this book focused on many approaches, seven are routinely used in data collection: surveys, tests, questionnaires, interviews, focus groups, observations, and performance records. The exercise that follows explores the use of each of these data collection methods.

---

**Instructions**

For each of the following situations, please indicate the most appropriate type of instrument to use in collecting the data needed to evaluate a program. Select from these types:

A. Survey
B. Test
C. Questionnaire
D. Interview
E. Focus group
F. Observation
G. Performance records

Write the letter of the most appropriate instrument in the box to the right of each question. Also indicate the level of evaluation to be pursued (1, 2, 3, or 4).

| *Situation* | *Instrument* | |
|---|---|---|
| 1. Customer service representatives have learned to resolve customer complaints in the most effective manner. An integral part of the program required customer service representatives to follow a series of planned steps to resolve complaints, using empathy and listening skills. As part of the evaluation, the human resources staff must determine the extent to which participants are actually using the newly acquired skills. | ☐ | ☐ |
| 2. Intact team members are involved in a conflict resolution program in which they acquired skills to resolve conflicts and disputes among themselves. Team members in this work group have a high degree of interaction and some of their responsibilities include checking the work of others. There had been an unusually high level of friction, with displays of open conflict in the group. In the program, participants learned how to deal with these issues and work together as a smooth operating team. The human resources staff needs to collect information about the group's progress, ideally in an environment in which there is an opportunity for | ☐ | ☐ |

group members to listen to comments from others.

3. Technicians participate in an e-learning program on basic mathematics and are required to achieve a predetermined level of competence in mathematics after completing the program. Human resources staff members measure the participants' level of mathematical ability before and after the program. ☐ ☐

4. The front desk staff at a major hotel have participated in a program to teach them how to use a new reservation system that is being installed. As part of the evaluation, it is important to obtain participants' reactions to the program and capture their planned actions. ☐ ☐

5. A company has implemented a new compensation plan in which the employees share in the overall profits of the company. Employees have attended a roll-out meeting in which they had the opportunity to learn how the program works and what is required of them to make it successful. As part of the evaluation, management is interested in finding out what the employees think about the new plan after attending the briefing. ☐ ☐

6. Sales representatives have a new commission program designed to improve sales. One objective of the program is to improve sales volume, and the human resources staff must determine exactly how much of an increase has been achieved by each individual since the program was conducted. ☐ ☐

7. Supervisors attended a problem-solving program in which they learned a logical ☐ ☐

approach to solving significant problems facing their work units. As a part of the program's evaluation, the human resources staff needs feedback from participants concerning their use of the acquired skills. The staff thinks there is a possibility that there is a success story here and will need to probe for details.

**Responses**

1. *Data collection method F, Level 3.* The key issue is the last sentence of the description. Observation is the only way to understand how someone is actually using these skills on the job. For the observation to work effectively, it must be invisible or unnoticeable. It would be Level 3 evaluation because data need to be collected about post-program skill usage.
2. *Data collection method E, Level 3.* The key to this issue is also in the last sentence. A focus group is the only method in which an opportunity to listen to the comments of others is available. This is important for capturing qualitative data. Data will be collected at Level 3 because the situation involves the use of skills.
3. *Data collection method B, Level 2.* This is a classic example of measuring learning, and a test is one of the most common ways of doing this. Testing for competencies is Level 2 data collection.
4. *Data collection method C, Level 1.* The desired instrument is considered a questionnaire because planned actions are to be captured. There is a clear distinction between surveys and questionnaires. A survey captures only attitudes, beliefs, and opinions, and responses are often solicited in an agree-or-disagree format or by means of a Likert scale representing the extent of agreement. A questionnaire may have survey data plus all other types of questions, including multiple-choice, yes-or-no, and fill-in-the-blank questions. This example captures Level 1 (reaction) data.

5. *Data collection method A, Level 1.* A survey can collect opinions about the participants' reactions to the plan. Gathering reactions is Level 1 evaluation.
6. *Data collection method G, Level 4.* This example requires data collection from record-keeping systems, which will provide credible hard data. Impact data (Level 4) are being collected in this example.
7. *Data collection method D, Level 3.* The interview is the only method that allows probing for more information in order to understand the responses or to explore an issue in more detail. In this case, Level 3 (application) data are collected.

---

## Selecting the Appropriate Method for Each Level

The data collection methods presented in this and earlier chapters offer a wide range of opportunities to collect data in a variety of situations. Nine issues should be considered when you are deciding on the most appropriate method of collecting data.

### Type of Data

One of the most important issues to consider when selecting a data collection method is the type of data to be collected. Some methods are more appropriate for collecting business impact data. Follow-up questionnaires, observations, interviews, focus groups, action planning, and performance contracting are best suited for application data; in fact, some of these methods are suited only for application data. Performance monitoring, action planning, and questionnaires can easily capture business impact data.

### Investment of Participants' Time

Another important factor in selecting a data collection method is the amount of time participants must spend with data collection and evaluation systems. Time requirements should always be

minimized, and the method should be positioned so that it is a value-added activity. Participants must understand that data collection is a valuable undertaking, not an activity to be resisted. Sampling can be helpful in keeping total time spent by participants to a minimum. Methods like performance monitoring require no participant time, whereas others, such as interviews and focus groups, require a significant investment of participants' time.

### Investment of Managers' Time

The time that a participant's manager must allocate to data collection is another issue in method selection. Time requirements for managers should always be minimized. Methods like performance contracting require significant involvement from the manager before and after program implementation, whereas other methods, such as participants' completion of a questionnaire, require no manager time.

### Cost

Cost is always a consideration in selecting the data collection method. Some data collection methods are more expensive than others. For example, interviews and observations are expensive, whereas surveys, questionnaires, and performance monitoring are usually inexpensive.

### Disruption of Normal Work Activities

Perhaps the issue that generates the greatest concern among managers is the degree of work disruption that data collection will create. Routine work processes should be disrupted as little as possible. Data collection techniques like performance monitoring require very little employee time and cause little distraction from normal activities. Questionnaires generally do not disrupt the work environment and can often be completed in just a few minutes, sometimes even after normal work hours. At the other extreme,

techniques such as focus groups and interviews may disrupt the work unit.

### Accuracy

The accuracy of the technique is another factor to consider in selecting a data collection method. Some data collection methods are more accurate than others. For example, performance monitoring is usually very accurate, whereas questionnaires are subject to distortion and may be unreliable, whether due to poor design or the mere subjectivity residing in individuals' responses. If on-the-job behavior must be captured, observation is clearly one of the most accurate methods. However, there is often a trade-off between the accuracy and the cost of a method.

### Built-In Design Possibility

Because building data collection into many evaluation plans is important, how easily the method can be built into a program is another consideration; it must become an integral part of the program. Some methods, such as action plans, can be easily built into the design of a program. Other methods, such as observation, are more difficult to integrate into a program.

In some situations, a program is redesigned to provide a follow-up session in which evaluation is addressed and additional training is offered. For example, an interactive selling skills program (a consecutive three-day program) was redesigned as a two-day workshop to build skills, followed by a one-day session three weeks later. The follow-up session provided an opportunity for additional training and evaluation. During the first part of the last day, Level 3 evaluation data were collected through a focus group. In addition, specific barriers and problems encountered when applying the skills were discussed. The second half of the day was devoted to additional skill building and refinement, along with techniques to overcome barriers to using the skills. The redesigned program provided a method for follow-up in both evaluation and training.

### Utility of an Additional Method

Because there are so many methods of collecting data, using too many methods is tempting. Multiple data collection methods add time and cost to an evaluation and may result in very little added value. Utility refers to the value added by each additional data collection method. When more than one method is used, this question should always be addressed: Does the value obtained from the additional data warrant the extra time and money that would need to spent on the method? If the answer is no, the additional method should not be implemented. The same issue must be addressed when considering multiple data sources and multiple time frames.

### Cultural Bias of Data Collection Method

The culture or philosophy of the organization may dictate which data collection methods would be best to use. For example, if an organization or audience is accustomed to using questionnaires, they will work well within the culture of that organization. If, however, an organization tends to overuse questionnaires, this may not be the best choice for collecting program data. Some organizations routinely use third-party observation. However, others view the technique as invasive—an attitude that is a clear deterrent to using observation to evaluate a program at that organization.

## Final Thoughts

This brief chapter has explored the issue of selecting an appropriate data collection method. In choosing a method, some straightforward criteria are useful. Most of the criteria detailed in this chapter will fit well in any organizational setting.

# Index

# About the Authors

**Patricia Pulliam Phillips, Ph.D.,** is president of the ROI Institute, Inc., the leading source of ROI competency building, implementation support, networking, and research. She supports organizations in their efforts to build accountability into their training, human resources, and performance improvement programs with a primary focus on building accountability in public sector organizations. She helps organizations implement the ROI Methodology in countries around the world, including South Africa, Singapore, Japan, New Zealand, Australia, Italy, Turkey, France, Germany, Canada, and the United States.

In 1997, after a thirteen-year career in the electrical utility industry, she embraced the ROI Methodology by committing herself to ongoing research and practice. To this end, Phillips has implemented the ROI Methodology in private sector and public sector organizations. She has conducted ROI impact studies of programs in leadership development, sales, new-hire orientation, human performance improvement, K–12 educator development, National Board Certification mentoring for educators, and faculty fellowship. Phillips is currently expanding her interest in public sector accountability by applying the ROI Methodology in community- and faith-based initiatives.

Phillips teaches others to implement the ROI Methodology through the ROI certification process, as a facilitator for ASTD's

ROI and Measuring and Evaluating Learning workshops, and as an adjunct professor for graduate-level evaluation courses. She speaks on the topic of ROI at conferences such as ASTD's International Conference and Exposition and the International Society for Performance Improvement's International Conference.

Phillips's academic accomplishments include a master's degree in public and private management and a Ph.D. degree in international development. She is certified in ROI evaluation and has earned the designation of certified performance technologist (CPT) and certified professional in learning and performance (CPLP). She has authored a number of publications on the subject of accountability and ROI, including *Show Me the Money: How to Determine ROI in People, Projects, and Programs* (Berrett-Koehler, 2007); *The Value of Learning* (Pfeiffer, 2007); *Return on Investment (ROI) Basics* (ASTD, 2005); *Proving the Value of HR: How and Why to Measure ROI* (Society for Human Resource Management, 2005); *Make Training Evaluation Work* (ASTD, 2004); *The Bottomline on ROI* (Center for Effective Performance, 2002), which won the 2003 ISPI Award of Excellence; *ROI at Work* (ASTD, 2005); the ASTD In Action casebooks *Measuring Return on Investment,* Volume 3 (2001), *Measuring ROI in the Public Sector* (2002), and *Retaining Your Best Employees* (2002); the ASTD Infoline series, including *Planning and Using Evaluation Data* (2003), *Mastering ROI* (1998), and *Managing Evaluation Shortcuts* (2001); and *The Human Resources Scorecard: Measuring Return on Investment* (Butterworth-Heinemann, 2001). Phillips's work has been published in a variety of journals. She can be reached at patti@roiinstitute.net.

---

**Cathy A. Stawarski, Ph.D.,** is program manager of the Strategic Performance Improvement and Evaluation program at the Human Resources Research Organization (HumRRO) in Alexandria, Virginia. She has more than twenty-five years experience in research, training and development, and program evaluation. Throughout

her nearly twenty years at HumRRO, she has worked primarily with clients in the federal sector. Her work includes leading and conducting the evaluation of leadership and human capital initiatives as well as assisting organizations in developing comprehensive evaluation strategies. Stawarski is a certified practitioner in the ROI Methodology and a member of the American Society for Training and Development and the International Society for Performance Improvement. She has authored nearly one hundred technical reports and training manuals during her career at HumRRO and has spoken at a variety of conferences on the topic of evaluation and ROI, including the Training Officers Conference, which supports the development of leaders in the federal government. She can be reached at cstawarski@humrro.org.

# Pfeiffer Publications Guide

This guide is designed to familiarize you with the various types of Pfeiffer publications. The formats section describes the various types of products that we publish; the methodologies section describes the many different ways that content might be provided within a product. We also provide a list of the topic areas in which we publish.

## FORMATS

In addition to its extensive book-publishing program, Pfeiffer offers content in an array of formats, from fieldbooks for the practitioner to complete, ready-to-use training packages that support group learning.

**FIELDBOOK** Designed to provide information and guidance to practitioners in the midst of action. Most fieldbooks are companions to another, sometimes earlier, work, from which its ideas are derived; the fieldbook makes practical what was theoretical in the original text. Fieldbooks can certainly be read from cover to cover. More likely, though, you'll find yourself bouncing around following a particular theme, or dipping in as the mood, and the situation, dictate.

**HANDBOOK** A contributed volume of work on a single topic, comprising an eclectic mix of ideas, case studies, and best practices sourced by practitioners and experts in the field.

An editor or team of editors usually is appointed to seek out contributors and to evaluate content for relevance to the topic. Think of a handbook not as a ready-to-eat meal, but as a cookbook of ingredients that enables you to create the most fitting experience for the occasion.

**RESOURCE** Materials designed to support group learning. They come in many forms: a complete, ready-to-use exercise (such as a game); a comprehensive resource on one topic (such as conflict management) containing a variety of methods and approaches; or a collection of like-minded activities (such as icebreakers) on multiple subjects and situations.

**TRAINING PACKAGE** An entire, ready-to-use learning program that focuses on a particular topic or skill. All packages comprise a guide for the facilitator/trainer and a workbook for the participants. Some packages are supported with additional media—such as video—or learning aids, instruments, or other devices to help participants understand concepts or practice and develop skills.

- *Facilitator/trainer's guide* Contains an introduction to the program, advice on how to organize and facilitate the learning event, and step-by-step instructor notes. The guide also contains copies of presentation materials—handouts, presentations, and overhead designs, for example—used in the program.
- *Participant's workbook* Contains exercises and reading materials that support the learning goal and serves as a valuable reference and support guide for participants in the weeks and months that follow the learning event. Typically, each participant will require his or her own workbook.

**ELECTRONIC** CD-ROMs and web-based products transform static Pfeiffer content into dynamic, interactive experiences. Designed to take advantage of the searchability, automation, and ease-of-use that technology provides, our e-products bring convenience and immediate accessibility to your workspace.

## METHODOLOGIES

**CASE STUDY** A presentation, in narrative form, of an actual event that has occurred inside an organization. Case studies are not prescriptive, nor are they used to prove a point; they are designed to develop critical analysis and decision-making skills. A case study has a specific time frame, specifies a sequence of events, is narrative in structure, and contains a plot structure—an issue (what should be/have been done?). Use case studies when the goal is to enable participants to apply previously learned theories to the circumstances in the case, decide what is pertinent, identify the real issues, decide what should have been done, and develop a plan of action.

**ENERGIZER** A short activity that develops readiness for the next session or learning event. Energizers are most commonly used after a break or lunch to

stimulate or refocus the group. Many involve some form of physical activity, so they are a useful way to counter post-lunch lethargy. Other uses include transitioning from one topic to another, where "mental" distancing is important.

**EXPERIENTIAL LEARNING ACTIVITY (ELA)** A facilitator-led intervention that moves participants through the learning cycle from experience to application (also known as a Structured Experience). ELAs are carefully thought-out designs in which there is a definite learning purpose and intended outcome. Each step—everything that participants do during the activity—facilitates the accomplishment of the stated goal. Each ELA includes complete instructions for facilitating the intervention and a clear statement of goals, suggested group size and timing, materials required, an explanation of the process, and, where appropriate, possible variations to the activity. (For more detail on Experiential Learning Activities, see the Introduction to the *Reference Guide to Handbooks and Annuals*, 1999 edition, Pfeiffer, San Francisco.)

**GAME** A group activity that has the purpose of fostering team spirit and togetherness in addition to the achievement of a pre-stated goal. Usually contrived—undertaking a desert expedition, for example—this type of learning method offers an engaging means for participants to demonstrate and practice business and interpersonal skills. Games are effective for team building and personal development mainly because the goal is subordinate to the process—the means through which participants reach decisions, collaborate, communicate, and generate trust and understanding. Games often engage teams in "friendly" competition.

**ICEBREAKER** A (usually) short activity designed to help participants overcome initial anxiety in a training session and/or to acquaint the participants with one another. An icebreaker can be a fun activity or can be tied to specific topics or training goals. While a useful tool in itself, the icebreaker comes into its own in situations where tension or resistance exists within a group.

**INSTRUMENT** A device used to assess, appraise, evaluate, describe, classify, and summarize various aspects of human behavior. The term used to describe an instrument depends primarily on its format and purpose. These terms include survey, questionnaire, inventory, diagnostic, survey, and poll. Some uses of instruments include providing instrumental feedback to group

members, studying here-and-now processes or functioning within a group, manipulating group composition, and evaluating outcomes of training and other interventions.

Instruments are popular in the training and HR field because, in general, more growth can occur if an individual is provided with a method for focusing specifically on his or her own behavior. Instruments also are used to obtain information that will serve as a basis for change and to assist in workforce planning efforts.

Paper-and-pencil tests still dominate the instrument landscape with a typical package comprising a facilitator's guide, which offers advice on administering the instrument and interpreting the collected data, and an initial set of instruments. Additional instruments are available separately. Pfeiffer, though, is investing heavily in e-instruments. Electronic instrumentation provides effortless distribution and, for larger groups particularly, offers advantages over paper-and-pencil tests in the time it takes to analyze data and provide feedback.

**LECTURETTE** A short talk that provides an explanation of a principle, model, or process that is pertinent to the participants' current learning needs. A lecturette is intended to establish a common language bond between the trainer and the participants by providing a mutual frame of reference. Use a lecturette as an introduction to a group activity or event, as an interjection during an event, or as a handout.

**MODEL** A graphic depiction of a system or process and the relationship among its elements. Models provide a frame of reference and something more tangible, and more easily remembered, than a verbal explanation. They also give participants something to "go on," enabling them to track their own progress as they experience the dynamics, processes, and relationships being depicted in the model.

**ROLE PLAY** A technique in which people assume a role in a situation/scenario: a customer service rep in an angry-customer exchange, for example. The way in which the role is approached is then discussed and feedback is offered. The role play is often repeated using a different approach and/or incorporating changes made based on feedback received. In other words, role playing is a spontaneous interaction involving realistic behavior under artificial (and safe) conditions.

**SIMULATION** A methodology for understanding the interrelationships among components of a system or process. Simulations differ from games in that they test or use a model that depicts or mirrors some aspect of reality in form, if not necessarily in content. Learning occurs by studying the effects of change on one or more factors of the model. Simulations are commonly used to test hypotheses about what happens in a system—often referred to as "what if?" analysis—or to examine best-case/worst-case scenarios.

**THEORY** A presentation of an idea from a conjectural perspective. Theories are useful because they encourage us to examine behavior and phenomena through a different lens.

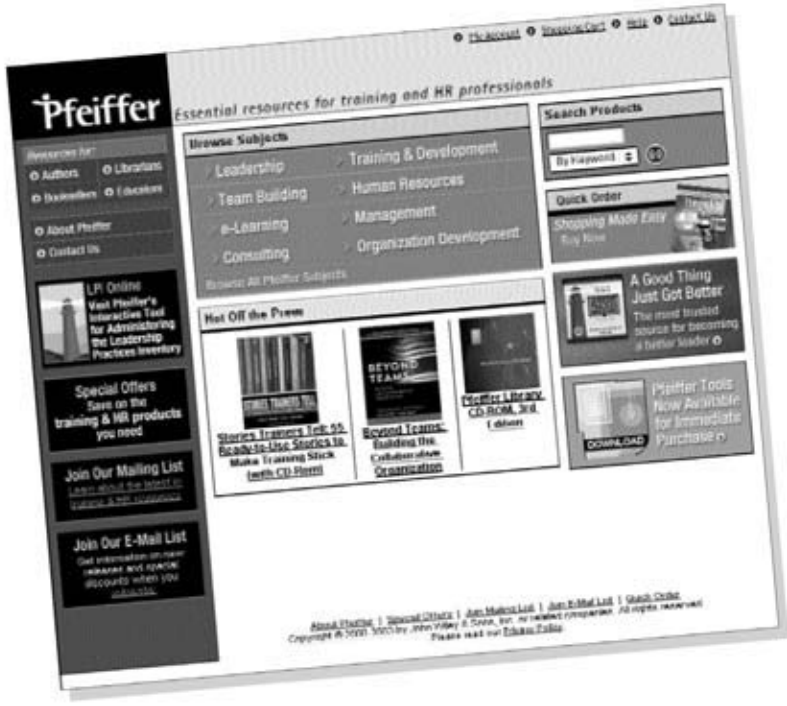
## TOPICS

The twin goals of providing effective and practical solutions for workforce training and organization development and meeting the educational needs of training and human resource professionals shape Pfeiffer's publishing program. Core topics include the following:

Leadership & Management
Communication & Presentation
Coaching & Mentoring
Training & Development
E-Learning
Teams & Collaboration
OD & Strategic Planning
Human Resources
Consulting

# What will you find on pfeiffer.com?

- The best in workplace performance solutions for training and HR professionals
- Downloadable training tools, exercises, and content
- Web-exclusive offers
- Training tips, articles, and news
- Seamless on-line ordering
- Author guidelines, information on becoming a Pfeiffer Affiliate, and much more




Discover more at www.pfeiffer.com

# Measurement and Evaluation Series

Series Editors
Patricia Pulliam Phillips, Ph.D., and Jack J. Phillips, Ph.D.

A six-book set that provides a step-by-step system for planning, measuring, calculating, and communicating evaluation and Return-on-Investment for training and development, featuring:

- Detailed templates
- Ready-to-use tools
- Complete plans
- Real-world case examples

The M&E Series features:

1. *ROI Fundamentals: Why and When to Measure ROI*
   (978-0-7879-8716-9)
2. *Data Collection: Planning For and Collecting All Types of Data*
   (978-0-7879-8718-3)
3. *Isolation of Results: Defining the Impact of the Program*
   (978-0-7879-8719-0)
4. *Data Conversion: Calculating the Monetary Benefits*
   (978-0-7879-8720-6)
5. *Costs and ROI: Evaluating at the Ultimate Level*
   (978-0-7879-8721-3)
6. *Communication and Implementation: Sustaining the Practice*
   (978-0-7879-8722-0)

Plus, the *ROI in Action Casebook* (978-0-7879-8717-6) covers all the major workplace learning and performance applications, including Leadership Development, Sales Training, Performance Improvement, Technical Skills Training, Information Technology Training, Orientation and OJT, and Supervisor Training.

The **ROI Methodology** is a comprehensive measurement and evaluation process that collects six types of measures: Reaction, Satisfaction, and Planned Action; Learning; Application and Implementation; Business Impact; Return on Investment; and Intangible Measures. The process provides a step-by-step system for evaluation and planning, data collection, data analysis, and reporting. It is appropriate for the measurement and evaluation of *all* kinds of performance improvement programs and activities, including training and development, learning, human resources, coaching, meetings and events, consulting, and project management.

## Special Offer from the ROI Institute

Send for your own ROI Process Model, an indispensable tool for implementing and presenting ROI in your organization. The ROI Institute is offering an exclusive gift to readers of The Measurement and Evaluation Series. This 11" × 25" multicolor foldout shows the ROI Methodology flow model and the key issues surrounding the implementation of the ROI Methodology. This easy-to-understand overview of the ROI Methodology has proven invaluable to countless professionals when implementing the ROI Methodology. Please return this page or e-mail your information to the address below to receive your free foldout (a $6.00 value). Please check your area(s) of interest in ROI.

Please send me the ROI Process Model described in the book. I am interested in learning more about the following ROI materials and services:

- ☐ Workshops and briefing on ROI
- ☐ Books and support materials on ROI
- ☐ Certification in the ROI Methodology
- ☐ ROI software
- ☐ ROI consulting services
- ☐ ROI Network information
- ☐ ROI benchmarking
- ☐ ROI research

Name _____

Title _____

Organization _____

Address _____

Phone _____

E-mail Address _____

### Functional area of interest:

- ☐ Learning and Development/Performance Improvement
- ☐ Human Resources/Human Capital
- ☐ Public Relations/Community Affairs/Government Relations
- ☐ Consulting
- ☐ Sales/Marketing
- ☐ Technology/IT Systems
- ☐ Project Management Solutions
- ☐ Quality/Six Sigma
- ☐ Operations/Methods/Engineering
- ☐ Research and Development/Innovations
- ☐ Finance/Compliance
- ☐ Logistics/Distribution/Supply Chain
- ☐ Public Policy Initiatives
- ☐ Social Programs
- ☐ Other (Please Specify) _____

### Organizational Level

☐ executive ☐ management ☐ consultant ☐ specialist
☐ student ☐ evaluator ☐ researcher

Return this form or contact The ROI Institute
P.O. Box 380637
Birmingham, AL 35238-0637

Or e-mail information to info@roiinstitute.net
Please allow four to six weeks for delivery.