



# A Guide to Evaluating Marine and Family Division Programs, Services, and Capabilities

Stephanie Brooks Holliday, Stephani L. Wrabel, Thomas E. Trail,  
Margaret Tankard, Dana Schultz, Laura Werber, Sarah O. Meadows



For more information on this publication, visit [www.rand.org/t/TL264](http://www.rand.org/t/TL264)

#### **About RAND**

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest. To learn more about RAND, visit [www.rand.org](http://www.rand.org).

#### **Research Integrity**

Our mission to help improve policy and decisionmaking through research and analysis is enabled through our core values of quality and objectivity and our unwavering commitment to the highest level of integrity and ethical behavior. To help ensure our research and analysis are rigorous, objective, and nonpartisan, we subject our research publications to a robust and exacting quality-assurance process; avoid both the appearance and reality of financial and other conflicts of interest through staff training, project screening, and a policy of mandatory disclosure; and pursue transparency in our research engagements through our commitment to the open publication of our research findings and recommendations, disclosure of the source of funding of published research, and policies to ensure intellectual independence. For more information, visit [www.rand.org/about/principles](http://www.rand.org/about/principles).

RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

Published by the RAND Corporation, Santa Monica, Calif.

© 2021 RAND Corporation

**RAND®** is a registered trademark.

Library of Congress Cataloging-in-Publication Data is available for this publication.

ISBN: 978-1-9774-0009-3

*Cover: Photo by U.S. Marine Corps Cpl. Alexander Mitchell.*

#### **Limited Print and Electronic Distribution Rights**

This document and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from RAND to reproduce, or reuse in another form, any of its research documents for commercial use. For information on reprint and linking permissions, please visit [www.rand.org/pubs/permissions](http://www.rand.org/pubs/permissions).

## Preface

---

Given increased emphasis on budgetary constraints and efficient spending, assessing not only the effectiveness of programs and services but also the efficiency of the overarching portfolio has gained significance. To date, the U.S. Marine Corps (USMC) Marine and Family Division (MF) has conducted numerous evaluations of different aspects of its portfolio of programs and services, including service quality studies, functionality assessments, industry benchmarking, compliance reviews, focus groups, and risk assessments. However, no overarching MF evaluation plan to assess either parts of the portfolio, in a comparable way, or the functioning of the portfolio as a whole currently exists. For MF programs and services to remain meaningful, credible, and practical, their effectiveness must be able to be demonstrated. This type of assessment will help the USMC to surge and contract the portfolio as resources wax and wane. Thus, the goal of this program evaluation toolkit is to help those responsible for implementing MF programs to determine whether those programs are effective—achieving their intended goals—and ultimately to guide efficient use of scarce resources.

This toolkit is composed of two interrelated documents: this User's Guide and a set of Excel workbooks that correspond to material presented in Chapter Six of the User's Guide. To develop the toolkit, the RAND team reviewed literature on program evaluation and program evaluation toolkits produced by other RAND researchers. Additional background for the specific context of MF, USMC, was provided by a review of relevant Marine Corps Orders (MCOs); discussion with headquarters, branch, and program staff in MF; and review of publicly available websites.

The contents of this toolkit should be of particular interest to MF and staff in the five branches within the scope of this study: Behavioral Health, Family Care, Family Readiness, Personal and Professional Development, and Semper Fit. It may also be of interest to those who manage or oversee similar programs that provide support to service members and their families.

This research was sponsored by the USMC MF and conducted within the Forces and Resources Policy Center of the RAND National Defense Research Institute, a federally funded research and development center sponsored by the Office of the Secretary of Defense, the Joint Staff, the Unified Combatant Commands, the Navy, the Marine Corps, the defense agencies, and the defense Intelligence Community.

For more information on the RAND Forces and Resources Policy Center, see <http://www.rand.org/nsrd/ndri/centers/frp.html> or contact the director (contact information is provided on the web page).



# Contents

---

|                                                                         |      |
|-------------------------------------------------------------------------|------|
| <b>Preface</b> .....                                                    | iii  |
| <b>Figures</b> .....                                                    | vii  |
| <b>Tables</b> .....                                                     | ix   |
| <b>Worksheets, Examples, and Info Boxes</b> .....                       | xi   |
| <b>Abbreviations</b> .....                                              | xiii |
| <br><b>CHAPTER ONE</b>                                                  |      |
| <b>Introduction and Toolkit Overview</b> .....                          | 1    |
| Intended Audience .....                                                 | 3    |
| How the Toolkit Was Developed .....                                     | 4    |
| Evaluation Challenges Addressed .....                                   | 4    |
| User's Guide .....                                                      | 5    |
| Summary .....                                                           | 6    |
| <br><b>CHAPTER TWO</b>                                                  |      |
| <b>Differentiating Among Programs, Services, and Capabilities</b> ..... | 7    |
| Program, Service, or Resource/Capability? .....                         | 7    |
| Summary .....                                                           | 11   |
| <br><b>CHAPTER THREE</b>                                                |      |
| <b>Identifying Core Components to Build a Logic Model</b> .....         | 13   |
| Identifying the Goals of Your Program, Service, or Capability .....     | 15   |
| Identifying Core Components .....                                       | 17   |
| Resources/Inputs .....                                                  | 18   |
| Activities .....                                                        | 21   |
| Target Population .....                                                 | 23   |
| Outputs .....                                                           | 25   |
| Outcomes .....                                                          | 26   |
| Impact .....                                                            | 29   |
| Context .....                                                           | 31   |
| Building the Logic Model .....                                          | 33   |
| Using the Logic Model for Your Evaluation .....                         | 38   |
| Additional Guidance .....                                               | 38   |
| Summary .....                                                           | 38   |

CHAPTER FOUR

|                                                                            |    |
|----------------------------------------------------------------------------|----|
| <b>Understanding Evaluation and Planning for Your Own Evaluation</b> ..... | 41 |
| Human Subjects Considerations .....                                        | 41 |
| Process Evaluation .....                                                   | 43 |
| Outcome Evaluation .....                                                   | 45 |
| Types of Evaluation Data .....                                             | 49 |
| Evaluation Considerations .....                                            | 52 |
| Selecting an Evaluation Design .....                                       | 59 |
| Summary .....                                                              | 71 |

CHAPTER FIVE

|                                                 |    |
|-------------------------------------------------|----|
| <b>Evaluation Measures</b> .....                | 75 |
| Measures of Performance and Effectiveness ..... | 75 |
| Defining Measures of Performance .....          | 75 |
| Defining Measures of Effectiveness .....        | 83 |
| Summary .....                                   | 91 |

CHAPTER SIX

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| <b>Conducting Your Evaluation</b> .....                           | 93  |
| Implementing an Evaluation Plan .....                             | 93  |
| Designing Processes to Collect Required Data .....                | 94  |
| Designing Tools to Collect Required Data .....                    | 97  |
| Analyzing Data .....                                              | 103 |
| Identifying Trends, Performance Outcomes, and Effectiveness ..... | 121 |
| Summary .....                                                     | 126 |

CHAPTER SEVEN

|                                             |     |
|---------------------------------------------|-----|
| <b>Continuous Quality Improvement</b> ..... | 127 |
| Plan-Do-Study-Act Cycle .....               | 128 |
| Conclusion .....                            | 134 |

|                         |     |
|-------------------------|-----|
| <b>References</b> ..... | 137 |
|-------------------------|-----|

## Figures

---

|       |                                                                                  |     |
|-------|----------------------------------------------------------------------------------|-----|
| 2.1.  | Identifying MF Offerings as PSCs.....                                            | 10  |
| 3.1.  | Logic Model Overview.....                                                        | 14  |
| 4.1.  | Framework for Evaluating a Program, Service, or Capability.....                  | 42  |
| 4.2.  | HRPP Applicability Review Process .....                                          | 44  |
| 5.1.  | Connection Between Logic Model and Evaluation Measures.....                      | 76  |
| 6.1.  | Sample Input to Use Conditional Formatting to Identify Values out of Range ..... | 98  |
| 6.2.  | Sample Output of Conditional Formatting to Identify Values out of Range .....    | 98  |
| 6.3.  | Sample Reverse-Scoring of a Scale Question .....                                 | 100 |
| 6.4.  | Sample Generation of an Average Score for a Scale.....                           | 101 |
| 6.5.  | Sample Use of a Group Membership Variable in Database.....                       | 102 |
| 6.6.  | Sample Use of Sorting by Group Membership Column.....                            | 102 |
| 6.7.  | Sample Use of Comments in Excel to Document Column Values.....                   | 103 |
| 6.8.  | Schematic of Workbooks in RAND Tool Based on Variable Type and Analysis.....     | 104 |
| 6.9.  | Entering and Coding Dichotomous Variables in RAND Tool Workbook.....             | 105 |
| 6.10. | Sample Histograms of Continuous Variables with Differing Distributions .....     | 106 |
| 6.11. | Pasting Continuous Indicators into RAND Tool Workbook.....                       | 107 |
| 6.12. | Sample Use of Formula to Count Number of Times a Value Appears.....              | 108 |
| 6.13. | Entering Data into a Time 1 Spreadsheet.....                                     | 110 |
| 6.14. | Sample Graph of a Dichotomous Indicator over Time.....                           | 111 |
| 6.15. | Sample Results of the Dichotomous Indicator Shown in Figure 6.14.....            | 112 |
| 6.16. | Write “Yes” or “No” in Yellow Cell for Tracking of Individuals.....              | 112 |
| 6.17. | Sample Graph of a Continuous Indicator over Time.....                            | 113 |
| 6.18. | Sample Results of the Continuous Indicator Shown in the Previous Graph .....     | 114 |
| 6.19. | Results of a Pre to Post Change in a Dichotomous Indicator .....                 | 115 |
| 6.20. | Notional Results of a Pre to Post Change in a Continuous Indicator.....          | 116 |
| 6.21. | Entering and Coding Group Data in the RAND Tool Workbook.....                    | 117 |
| 6.22. | The Group Difference Results Spreadsheet .....                                   | 117 |
| 6.23. | Entering Group Data into the RAND Tool Spreadsheet.....                          | 118 |
| 6.24. | Group Difference Results Spreadsheet .....                                       | 118 |
| 6.25. | Correlation Workbook Showing Moderately Correlated Variables.....                | 120 |
| 7.1.  | Plan-Do-Study-Act Cycle .....                                                    | 127 |
| 7.2.  | Identifying Approaches for Improvement.....                                      | 129 |



## Tables

---

|      |                                                                                     |     |
|------|-------------------------------------------------------------------------------------|-----|
| 3.1. | Determining Whether Goals Meet the SMART Principles.....                            | 16  |
| 4.1. | Questions That Guide Process Evaluation.....                                        | 46  |
| 4.2. | Advantages and Disadvantages of Outcome Evaluation Design Options .....             | 50  |
| 4.3. | Examples of Qualitative, Quantitative, Objective, and Subjective Data.....          | 52  |
| 5.1. | Characteristics of Measures of Performance and Measures of Effectiveness .....      | 77  |
| 5.2. | Sample MOPs for Evaluating Program Participation .....                              | 77  |
| 5.3. | L.I.N.K.S. Program MOPs.....                                                        | 80  |
| 5.4. | FMEAP Service MOPs .....                                                            | 81  |
| 5.5. | Outdoor Recreation Equipment Checkout and Equipment Rental<br>Capability MOPs ..... | 82  |
| 5.6. | Sample L.I.N.K.S. Program MOEs .....                                                | 87  |
| 5.7. | Sample FMEAP Service MOEs.....                                                      | 88  |
| 5.8. | Outdoor Recreation Equipment Checkout and Equipment Rental<br>Capability MOEs ..... | 89  |
| 6.1. | Sample Sources of Misinterpretation.....                                            | 124 |
| 7.1. | Tracking Areas Identified for Program Improvement .....                             | 130 |



## Worksheets, Examples, and Info Boxes

---

### **Info Box 1.1.**

|                                                      |   |
|------------------------------------------------------|---|
| Program Evaluation Versus Portfolio Management ..... | 2 |
|------------------------------------------------------|---|

### **Worksheet 3.1.**

|                               |    |
|-------------------------------|----|
| Identifying SMART Goals ..... | 17 |
|-------------------------------|----|

### **Worksheet 3.2.**

|                                                                                                             |    |
|-------------------------------------------------------------------------------------------------------------|----|
| Resources/Inputs .....                                                                                      | 18 |
| Example Resources/Inputs for L.I.N.K.S. Program .....                                                       | 19 |
| Example Resources/Inputs for FMEAP Service .....                                                            | 19 |
| Example Resources/Inputs for Outdoor Recreation Equipment Checkout and<br>Equipment Rental Capability ..... | 19 |

### **Worksheet 3.3.**

|                                                                                                     |    |
|-----------------------------------------------------------------------------------------------------|----|
| Activities .....                                                                                    | 20 |
| Example Activity for L.I.N.K.S. Program .....                                                       | 21 |
| Example Activity for FMEAP Service .....                                                            | 21 |
| Example Activity for Outdoor Recreation Equipment Checkout and<br>Equipment Rental Capability ..... | 21 |

### **Worksheet 3.4.**

|                                                                                                     |    |
|-----------------------------------------------------------------------------------------------------|----|
| Target Population .....                                                                             | 22 |
| Example Target Population for L.I.N.K.S. Program .....                                              | 23 |
| Example Target Population for FMEAP Service .....                                                   | 23 |
| Example Activity for Outdoor Recreation Equipment Checkout and<br>Equipment Rental Capability ..... | 23 |

### **Worksheet 3.5.**

|                                                                                                    |    |
|----------------------------------------------------------------------------------------------------|----|
| Outputs .....                                                                                      | 24 |
| Example Outputs for L.I.N.K.S. Program .....                                                       | 25 |
| Example Outputs for FMEAP Service .....                                                            | 25 |
| Example Outputs for Outdoor Recreation Equipment Checkout and<br>Equipment Rental Capability ..... | 25 |

### **Worksheet 3.6.**

|                                                                                                     |    |
|-----------------------------------------------------------------------------------------------------|----|
| Outcomes .....                                                                                      | 26 |
| Example Outcomes for L.I.N.K.S. Program .....                                                       | 27 |
| Example Outcomes for FMEAP Service .....                                                            | 27 |
| Example Outcomes for Outdoor Recreation Equipment Checkout and<br>Equipment Rental Capability ..... | 28 |

### **Worksheet 3.7.**

|                                             |    |
|---------------------------------------------|----|
| Impact .....                                | 29 |
| Example Impact for L.I.N.K.S. Program ..... | 30 |

|                                                                                                        |     |
|--------------------------------------------------------------------------------------------------------|-----|
| Example Impact for FMEAP Service.....                                                                  | 30  |
| Example Impact for Outdoor Recreation Equipment Checkout and<br>Equipment Rental Capability .....      | 30  |
| <b>Worksheet 3.8.</b>                                                                                  |     |
| Context .....                                                                                          | 31  |
| Example Context for L.I.N.K.S. Program.....                                                            | 32  |
| Example Context for FMEAP Service .....                                                                | 32  |
| Example Context for Outdoor Recreation Equipment Checkout and<br>Equipment Rental Capability .....     | 32  |
| <b>Worksheet 3.9.</b>                                                                                  |     |
| The Logic Model .....                                                                                  | 34  |
| Example Logic Model for L.I.N.K.S. Program .....                                                       | 35  |
| Example Logic Model for FMEAP Service.....                                                             | 36  |
| Example Logic Model for Outdoor Recreation Equipment Checkout and<br>Equipment Rental Capability ..... | 37  |
| <b>Worksheet 4.1.</b>                                                                                  |     |
| Identifying the Focus of Your Process Evaluation .....                                                 | 47  |
| Example Identifying the Focus of Your Process Evaluation .....                                         | 48  |
| <b>Worksheet 4.2.</b>                                                                                  |     |
| Evaluation Design Issues.....                                                                          | 60  |
| <b>Worksheet 4.3.</b>                                                                                  |     |
| Designing Your Evaluation .....                                                                        | 62  |
| Example L.I.N.K.S. Program Evaluation Plan.....                                                        | 64  |
| Example FMEAP Service Evaluation Plan .....                                                            | 66  |
| Example Outdoor Recreation Equipment Checkout and Equipment Rental Capability<br>Evaluation Plan ..... | 69  |
| <b>Worksheet 4.4.</b>                                                                                  |     |
| Evaluation Gantt Chart.....                                                                            | 72  |
| Example GANTT Chart for L.I.N.K.S. Program.....                                                        | 73  |
| <b>Worksheet 5.1.</b>                                                                                  |     |
| Operationalizing Measures of Performance .....                                                         | 83  |
| <b>Worksheet 5.2.</b>                                                                                  |     |
| Operationalizing Measures of Effectiveness .....                                                       | 90  |
| <b>Worksheet 6.1.</b>                                                                                  |     |
| Review Program Outcomes.....                                                                           | 122 |
| Example Review Program Outcomes Using L.I.N.K.S. ....                                                  | 123 |
| <b>Worksheet 7.1.</b>                                                                                  |     |
| Tracking Areas Identified for Program Improvement .....                                                | 131 |
| <b>Info Box 7.1.</b>                                                                                   |     |
| Program Evaluation and Strategic Portfolio Management (SPM) .....                                      | 134 |

## Abbreviations

---

|            |                                                            |
|------------|------------------------------------------------------------|
| CQI        | Continuous Quality Improvement                             |
| DoD        | U.S. Department of Defense                                 |
| DoDI       | Department of Defense Instruction                          |
| FMEAP      | Family Member Employment Assistance Program                |
| HRPP       | Human Subjects Research Protection Program                 |
| L.I.N.K.S. | Lifestyle Insights, Networking, Knowledge, and Skills      |
| M4L        | Marine for Life                                            |
| MCCS       | Marine Corps Community Services                            |
| MCO        | Marine Corps Order                                         |
| MF         | Marine and Family Division                                 |
| MOE        | measures of effectiveness                                  |
| MOP        | measures of performance                                    |
| PDSA       | Plan-Do-Study Act                                          |
| PSC        | program, service, or capability                            |
| RCT        | randomized controlled trial                                |
| SMART      | Specific, Measurable, Achievable, Relevant, and Time-bound |
| SPM        | strategic portfolio management                             |
| USMC       | U.S. Marine Corps                                          |



## Introduction and Toolkit Overview

---

In an effort to help service members and their families navigate the stresses associated with deployments and other aspects of military life, the U.S. Marine Corps (USMC) has developed a set of programs, services, and policies that provide support to Marines and their families. The Marine and Family Division (MF) portfolio includes a variety of programs and services that cut across a number of domains, including behavioral health, personal and professional development, and family readiness. Because these resources are associated with Marine personnel readiness, the portfolio is indirectly, and critically, tied to the USMC mission and its future success.

The eight programmatic branches of the MF portfolio include

- Behavioral Health Services
- Sexual Assault Prevention and Response
- Family Care
- Family Readiness
- Personal and Professional Development
- Wounded Warrior Regiment
- Military Personnel Services
- Semper Fit.

Given increased emphasis on budgetary constraints and efficient spending, assessing not only the effectiveness of programs and services but also the efficiency of the overarching portfolio has gained significance. To date, MF has conducted numerous evaluations of different aspects of the portfolio, including service quality studies, functionality assessments, industry benchmarking, compliance reviews, focus groups, and risk assessments. However, no overarching MF evaluation plan to assess either parts of the portfolio, in a comparable way, or the functioning of the portfolio as a whole exists. For MF programs and services to remain meaningful, credible, and practical, their effectiveness must be able to be demonstrated. This type of assessment will help USMC to surge and contract the portfolio as resources wax and wane.

Program evaluation is the application of systematic processes for collecting and analyzing information about a program to determine that program's quality of operation and ability to meet intended purposes (i.e., effectiveness) (Rossi, Lipsey, and Freeman, 2004; Wholey, Hatry, and Newcomer, 2010). At its core, program evaluation is used to guide decisionmaking about the right set of activities to offer; inform efforts to improve current activities; identify areas of practice that are at, above, or below expectation; and demonstrate the value of a particular program in effectively contributing to high-priority goals within an organization.

**Info Box 1.1****Program Evaluation Versus Portfolio Management**

Program evaluation is an aspect of strategic portfolio management (SPM). Each MF branch manages a portfolio of programs meant to work collectively to meet the broad purposes of that branch, such as increasing retention or improving resiliency of active-duty Marines. SPM outlines a process that guides the portfolio design, provides ongoing management and evaluation of the portfolio, and helps to determine resource allocation across a set of programs. Portfolio evaluation can be used to assess whether the portfolio is being well managed, whether it contains programs that are meaningfully connected, and whether the portfolio—and its programs—are agile. (For more on SPM, see Info Box 7.1.)

This toolkit will provide more information about SPM in Chapter Seven.

Program evaluation is not valuable as a tool only for programs that appear to be struggling to meet intended purposes. Even well-implemented and seemingly effective programs can use the results of an evaluation to improve the efficient allocation of resources, streamline the delivery of program components, and turn good programs into great programs for members of our military services. Moreover, these programs will eventually experience changes in staff, introduction of new policy or mandates, or reductions or additions to operating budgets. In such cases, these previously effective programs may need to adjust operations in ways that will influence a program's ability to meet the established goals (Ryan et al., 2014). Program evaluation can help guide adjustments in well-established practices.

The broader purpose of program evaluation is to assist program managers in cycles of continuous program improvement, though program evaluations can be conducted as one-time events. We discuss more about continuous and one-time use evaluations in Chapter Seven, when we review the concept of Continuous Quality Improvement (CQI). Whether for continuous or one-time evaluations, the planning and preparation for an evaluation cycle needs to be an intentional activity at the beginning of a program cycle. Waiting until after program operations have concluded limits both what can be learned and the appropriateness of the results to address all aspects of program operations. Moreover, having an established and written evaluation plan can foster transparency for stakeholders and decisionmakers, increase buy-in and support for the evaluation efforts, or serve as an advocacy tool for gaining resources needed to complete the evaluation (National Center for Chronic Disease Prevention and Health Promotion, 2011).

The overall goal of this toolkit is to help USMC MF program managers and staff to both (1) understand the components and processes of program evaluation, and (2) successfully prepare and conduct rigorous evaluations of their own programs. Ultimately, this toolkit is a guide for starting a cycle of continuous improvement conducted by each offering within the MF portfolio. It contains two interrelated products: a User's Guide (this document) and a set of Microsoft Excel files that help facilitate analyses.

To accomplish these goals, this toolkit aims to

- provide an understanding of the theory and foundation of evaluations, which will help guide all evaluation and continuous improvement efforts
- help program managers design an evaluation that is appropriate given the nature of the program, as well as the resources and expertise available for the evaluation
- discuss ways to select and operationalize evaluation measures
- offer guidance on how to analyze evaluation data, including instructions on how to use a data analytic tool specifically aligned to this toolkit

- explain how results from the evaluation can then be used to improve the operation and effectiveness of the program
- review the practices of a CQI cycle to establish a culture of continuous improvement.

Each chapter of the User's Guide contributes to these specific aims, which can be achieved by working through the entire document. Chapter Two presents a framework to help explain differences between MF programs, services, and resources and capabilities, which have implications for the types of information collected and analyzed during an evaluation. We identify specific MF offerings in each of these categories, which are used to demonstrate the program evaluation process throughout the remainder of the User's Guide. Chapter Three enables you to identify a program's core components and how these components are interconnected, which will serve as a foundation for designing its evaluation. Chapters Four and Five, respectively, describe specific evaluation designs and measures. Chapter Four guides the user to weigh the benefits and costs of specific evaluation designs in order to select an evaluation design appropriate for a program's resources, expertise, and needs. In Chapter Five, the user learns to select and define relevant evaluation measures. Chapter Six discusses the process of collecting and analyzing evaluation data, with specific guidance on how to use the Analytic Tool (i.e., Excel workbooks) developed to accompany this toolkit. Finally, Chapter Seven discusses the use of evaluation data for CQI.

Before we proceed to Chapter Two of the User's Guide, we review the intended audience of this toolkit, present how the toolkit was developed, detail how program evaluation is one component of strategic portfolio management, and identify some of the key issues related to evaluations that we discuss throughout the remaining chapters. We conclude with a brief overview of the User's Guide content and how we recommend using the tools presented throughout.

## Intended Audience

This toolkit has been developed to be used primarily by program staff implementing one of the programs within the MF branches. The guidance and tools presented in this toolkit have been developed with a focus on five of the eight programmatic branches: Behavioral Health, Family Care, Family Readiness, Personal and Professional Development, and Semper Fit. The information presented in this toolkit may also serve as a resource for program staff across the MF branches or for the sister military services executing similar programs that provide support to service members and their families.

Regardless of who finds this toolkit helpful, all users must ensure that they are knowledgeable and respectful of the policy, standards, and guidance established by the Department of Defense (DoD) and the USMC for evaluations and research that involve human subjects (see Chapter Four of this toolkit for more details).

For individuals who may want a more condensed overview, a brief introduction to program evaluation can be found in *A Guide to Evaluating Marine and Family Division Programs, Services, and Capabilities: An Introduction to Program Evaluation* (Holliday et al., 2021).

## How the Toolkit Was Developed

To create this toolkit (both the User's Guide and the associated workbook files), we conducted a review of the program evaluation literature, including evaluation studies of family resilience and support programs and existing evaluation toolkits. The literature review was used to ensure the evaluation processes presented in this toolkit align with the latest research on evaluation practices and to identify measures that can be used to assess both how programs are implemented and their effectiveness. The review of existing toolkits informed the outline of this toolkit's components and our decisions about the types of tools that are included (e.g., workbooks). Specifically, we relied heavily on prior RAND work, including

- *RAND Suicide Prevention Program Evaluation Toolkit* (Acosta et al., 2013),
- *Programs Addressing Psychological Health and Traumatic Brain Injury Among U.S. Military Servicemembers and Their Families* (Weinick et al., 2011).

To ensure that we had a comprehensive understanding of each of the MF programs, we reviewed all current Marine Corps Orders (MCOs) for these programs, searched for public documentation, and conducted interviews with a subset of MF branch managers and program managers. The MCOs and other program documentation covered important details, such as the activities offered, processes for utilizing a service or enrolling in a program, and how the programs are marketed to target populations. Finally, the interviewees offered insights about current program operations, previous or current evaluation and assessment efforts, and external factors affecting the programs at a systemwide level.

Because the branches within MF cover multiple professional disciplines and perspectives, we have worked to find common ground in the language used throughout this toolkit rather than tailoring the language to meet any one branch or discipline's preferred terminology. In this way, the language used in the toolkit may vary somewhat from the terms you are familiar with and use on a regular basis. The terminology used here also reflects industry standards common to program evaluation, so that users can find additional resources and references with relative ease. Finally, the common language also supports strategic portfolio management (see Info Box 7.1), as all MF programs are intended to work collaboratively to serve Marines and their families. Each MF branch functions at the installation or force level, moving beyond the scope of work served by each individual program at an installation. The common language allows for establishing connection and discussions between programs, services, or capabilities (PSCs) to examine the entire MF portfolio.

## Evaluation Challenges Addressed

Program staff may face a few challenges when evaluating programs within MF. We summarize these challenges next and provide methods to address each of them throughout the toolkit. It is important to keep these challenges in mind while using the tools presented in this toolkit.

**Challenge One: The types of programmatic offerings that make up the whole MF portfolio are diverse in their purposes, operations, target populations, and resource levels.**

In Chapter Two of this toolkit, we distinguish between the different types of programs to help the user tailor evaluation components and processes to meet a specific program's needs. We revisit these three program types (programs, services, and capabilities) in the subsequent chapters, calling out ways in which evaluation methods may differ for each of the program types and providing guidance wherever possible.

**Challenge Two: Within individual programs, managers must account for multiple activities and programmatic responsibilities at the same time, making it difficult to isolate the components or characteristics of that program that are uniquely responsible for any observed effect.**

The toolkit supports the development of a program logic model (described in Chapter Three) that ties specific program activities to specific short-term and intermediate-term outcomes. Identifying which outcomes have been achieved can help program staff better understand the drivers of any changes, such as increased unit cohesion. In the final chapter of this toolkit, we provide some guidance on how program managers can consider modifying aspects of their programs to address areas of need when some of the desired outcomes are unmet or a program has not been implemented as intended.

**Challenge Three: Some programs contend with confidentiality issues, have anonymous participants, or contract out services to external providers.**

Each of these operating practices presents unique challenges for the staff charged with demonstrating program effectiveness. This toolkit reviews some of the considerations associated with each of these issues to ensure a rigorous, appropriate evaluation can be conducted.

## User's Guide

### Overview of Content

The remainder of this document provides an overview of critical program evaluation practices and guides the user through a series of worksheets. These worksheets will help you

- create a detailed and complete logic model that serves as the foundation for the program evaluation design (Chapter Three)
- design your program evaluation, including the questions that guide the evaluation, sources of data, and timeline for the evaluation (Chapter Four)
- specify process and outcome measures (Chapter Five).

Finally, Chapter Six provides an overview of basic data collection and analytic practices and demonstrates how to use the sister analysis tool for such analyses. A set of Excel workbooks is available to help the user practice these exercises.

As described previously, this toolkit concludes with an overview of a continuous quality improvement cycle and how incorporating these cycles into routine practice helps improve the program provided to military service members and their families (Chapter Seven).

This toolkit is most effectively used by working through each step sequentially. However, for programs that have previously conducted an evaluation, or whose managers have existing evaluation expertise, it is possible to use these tools on a more ad hoc or modular basis.

## **Summary**

The reader should now be familiar with the purpose of program evaluation, the target audience for this particular toolkit, and how it was developed. We also introduced the concepts that will be presented in each of the following chapters. This User's Guide is designed such that a user can proceed sequentially through each chapter, completing worksheets and tasks throughout, to complete an evaluation. In the next chapter, we will distinguish between the three types of MF offerings—which we refer to as programs, services, and capabilities—that the USMC provides for active-duty personnel, reservists, family members, and (on occasion) veterans and civilians. Such distinctions will help a team plan an appropriate and rigorous evaluation.

## Differentiating Between Programs, Services, and Capabilities

---

Up to this point in the User's Guide, we have largely used the term *program* to refer to all of the MF offerings. Generally speaking, a *program* is any set of coordinated activities or resources conducted to reach a goal or fulfill a particular purpose (Wholey, Hatry, and Newcomer, 2010). This broad definition holds true across all of the MF offerings. However, for the purposes of this User's Guide, we make an additional distinction between *programs* (P), *services* (S), and *resources/capabilities* (C) and refer to the complete set of what MF provides as *PSCs*. This distinction has implications for understanding how a program is expected to operate, the purposes of the program, and what an evaluation can be expected to demonstrate about that program. In this chapter, we describe the features of programs, services, and capabilities, which will enable you and your team to situate your own PSC. The discussion of differences between PSCs will continue throughout the toolkit to ensure the evaluation your team designs is appropriate for your PSC type.

### Program, Service, or Resource/Capability?

The Innovative Practices project (Weinick et al., 2011), a prior study at RAND examining psychological health programs and services for members of the military community, distinguished between different types of programs based on level of interaction between providers and participants, whether activities were part of routine operations versus specialized activities, and clinical versus nonclinical nature of the services offered. We have adapted the framework presented by Innovative Practices to capture the variation in PSCs that make up each branch and the larger MF portfolio.

Admittedly, the distinctions between programs, services, and capabilities can be difficult to identify, especially when comparing services and capabilities. This is made more difficult in that MF also uses the language of programs and services in its distinction of portfolio offerings. To be clear, the terminology established by Weinick and colleagues (2011) around PSCs was developed for evaluation purposes and is not meant to supersede, replace, or change how MF identifies its offerings as programs or services. Moreover, the PSC terminology used in this toolkit does not reflect a hierarchy of importance or value to the overall MF portfolio.

We further explain key features of PSCs next and provide examples directly from the MF portfolio. Importantly, the PSC distinction used in this toolkit does not shape how an evaluation is conducted but is intended to assist program managers in understanding what an evaluation can tell them about their operations and effectiveness if operating a program, service, or capability. As discussed more throughout the toolkit, programs may be able to generate

changes in knowledge, skills, behaviors, or attitudes, while capabilities may increase access or connection to resources. The toolkit is created in such a way that no matter what your team operates, it will still help guide the evaluation of what is provided to Marines and their family members.

**Programs** provide activities, interventions, or other efforts to meet an overall goal or purpose. Above all, programs are interactive in nature, as participants engage with trained program personnel, staff, or other active-duty personnel or family members in structured ways. Participants commonly must register or qualify to participate, as programs are not carried out as day-to-day activities. Programs may include a capability component (described below), but they must go beyond passive dissemination of information.

**Specific MF Example:** *Lifestyle Insights, Networking, Knowledge, and Skills (L.I.N.K.S.).* L.I.N.K.S. is an interactive acculturation program in which participants attend sessions to learn about USMC history and local installation resources and services, as well as terminology and key aspects of military life. These acculturation sessions are differentiated for active-duty members and dependents (i.e., spouses, teens, and children), as well as parents and extended family members to ensure the content is developmentally appropriate. Overall, the purpose of L.I.N.K.S. is to ensure mission readiness by educating and preparing Marines and their family members to successfully navigate and understand military life.

**Services** are part of what the USMC routinely provides to service members and their families, and, to a lesser extent, civilian personnel, retirees, and contractors. Services can contribute to important outcomes like quality of life, satisfaction with military life, and ultimately retention. Some services may offer programming for active-duty personnel and their dependents that includes active engagement with staff or other members of the military community. However, such programming is a practice used to reach specific intended goals and is not the primary method of operation.

**Specific MF Example:** *Family Member Employment Assistance Program (FMEAP).* This service provides employment-related services to active-duty personnel, the spouses and dependents of active-duty military personnel, retirees, and DoD civilians. FMEAP offices provide access to computers, fax machines, and copiers; maintain resource libraries that include local labor market and occupational information; and offer personal assessments, job search training, and resume or job application assistance. The overall purpose of this service is to empower military-connected persons with the necessary tools to achieve their career, education, or volunteering goals.

**Capabilities** are structured as a one-way, passive transmission of information, such as an information brochure, website, or directory that lists services available on a specific installation. Some capabilities may operate in ways similar to services, by engaging with members of the military community when staff members are sought out for additional information or personalized assistance, but engagement is not a critical component of what occurs. In many cases, capabilities are operations that support other programs or services within the portfolio of offerings operated by MF program division. For example, Family Readiness Program Training helps to support

the Family Readiness Officer service. However, the key difference is that the capability is not directly responsible for the everyday operation of those programs or services.

**Specific MF Example:** *Outdoor Recreation Equipment Checkout and Equipment Rental.* The focus of Outdoor Recreation Equipment Checkout and Equipment Rental is to provide equipment support for directed and self-directed participation in outdoor activities, resale, and equipment rental services. This capability provides individual Marines and their dependents access to low-cost recreational equipment for use during off-duty hours, supports programs offered by MF (e.g., Directed Outdoor Recreation) by providing necessary supplies and equipment, and fulfills equipment requests made by units and commands for local or large-scale events. While staff may offer insight or instruction on how to use a particular piece of equipment or be active in the delivery of equipment, this operation is predominantly passive in its provision of support to Marines and their dependents. In fact, many installations maintain a website, in addition to the physical Equipment Rental location, where available supplies can be identified and requested.

We highlight these examples of PSCs to assist in your understanding of how to differentiate between each type of offering. We will continue to use these three PSCs to demonstrate the concepts and activities presented in the remainder of the toolkit.

Describing each of these also shows some of the fuzziness in distinguishing between the three PSC types within the MF portfolio. In fact, some services may have components that resemble capabilities (e.g., maintaining a website for individuals to seek out for information), and some services may have components that reflect programs (e.g., a service offering a program about proper nutrition for the physical demands of Marine Corps life). We used the core mission and operations to guide the identification of an offering in the portfolio as a program, service, or capability.

In Figure 2.1, we provide a flowchart of questions to assist program and branch managers in identifying whether a specific offering is a program, service, or capability for the purposes of evaluation. Again, the distinctions among these categories can be fuzzy. When you consider each question, ask yourself what are the **primary** activities in the offering. For example, if an offering provides referrals and information 85 percent of the time, but also has a small number of interactive workshops, your conclusion would be that your offering primarily involves passive or one-way transmission of information. When in doubt about which type of offering your PSC falls into, we recommend selecting the more active option (e.g., a service when debating between service and capability) to guide your evaluation work.

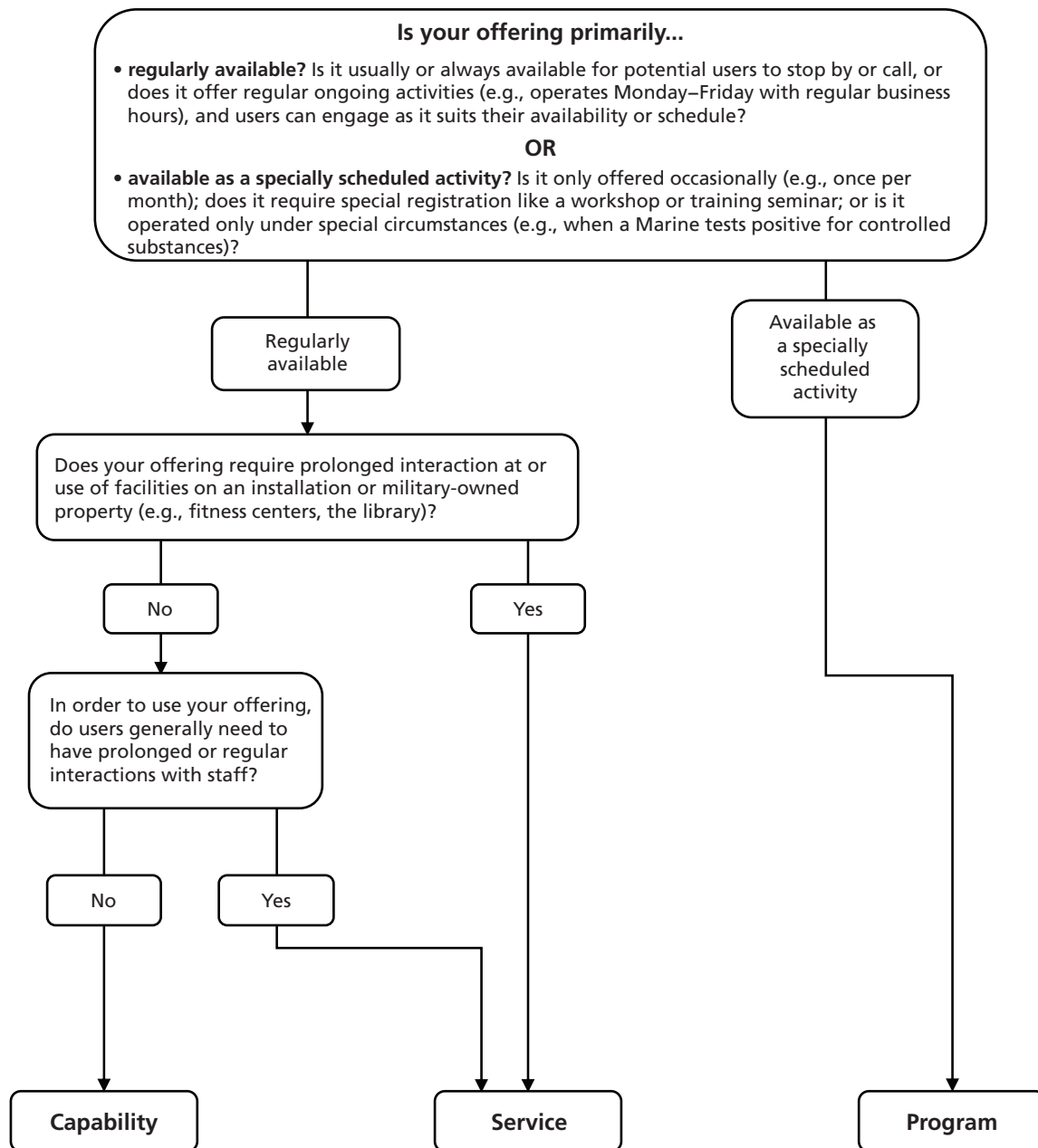
The flowchart in Figure 2.1 can be applied to our sample program, service, and capability. For example, Outdoor Recreation Equipment Checkout and Equipment Rental (a) is routinely available, (b) makes use of Marine Corps equipment and/or property, and (c) does not require sustained or regular interactions with staff. Therefore, it is appropriately classified as a capability.

This flowchart can also be used with less clear-cut examples. For instance, Marine for Life (M4L) is an MF offering that helps to connect transitioning Marines and their family members to resources that will aid their transition to civilian life. In addition to providing referrals to local resources, M4L holds local, in-person networking events for separated Marines. Although these networking events might be considered “specially scheduled activities,” they are not the primary activity of M4L. Most of M4L’s activities are available on an ongoing basis,

including a Facebook and LinkedIn community for virtual networking. Therefore, this would be classified as routinely available. Because the M4L activities do not take place on an installation or military-owned property, and do not require sustained or regular interactions with staff, it would be best classified as a capability under this framework.

We continue to reference the PSCs used in this chapter throughout the toolkit to show where differences between PSCs affect how an organization designs and conducts an appropriate, rigorous evaluation. As mentioned above, though, no matter what types of activities your

**Figure 2.1**  
**Identifying MF Offerings as PSCs**



PSC offers, this toolkit should still be useful in designing and implementing a tailored evaluation plan.

## **Summary**

In this chapter, we distinguished between the three types of offerings (i.e., programs, services, and capabilities) that the MF provides for military personnel and their families and on a limited basis for civilian personnel, retirees, and contractors. Identifying the main purposes and functioning of what your team offers will assist in guiding you through this toolkit and the evaluation process. This starts in the next chapter, with our discussion of building a logic model, which reflects how your program, service, or capability is expected to meet its purpose and goals.



## Identifying Core Components to Build a Logic Model

---

In this chapter, we will help you identify the core components of your program, service, or capability, including the available resources, activities, target population, and intended outcomes and impact. Then, we will guide you through a process for organizing these elements into a logic model that can be the basis for an evaluation plan to assess performance and effectiveness.

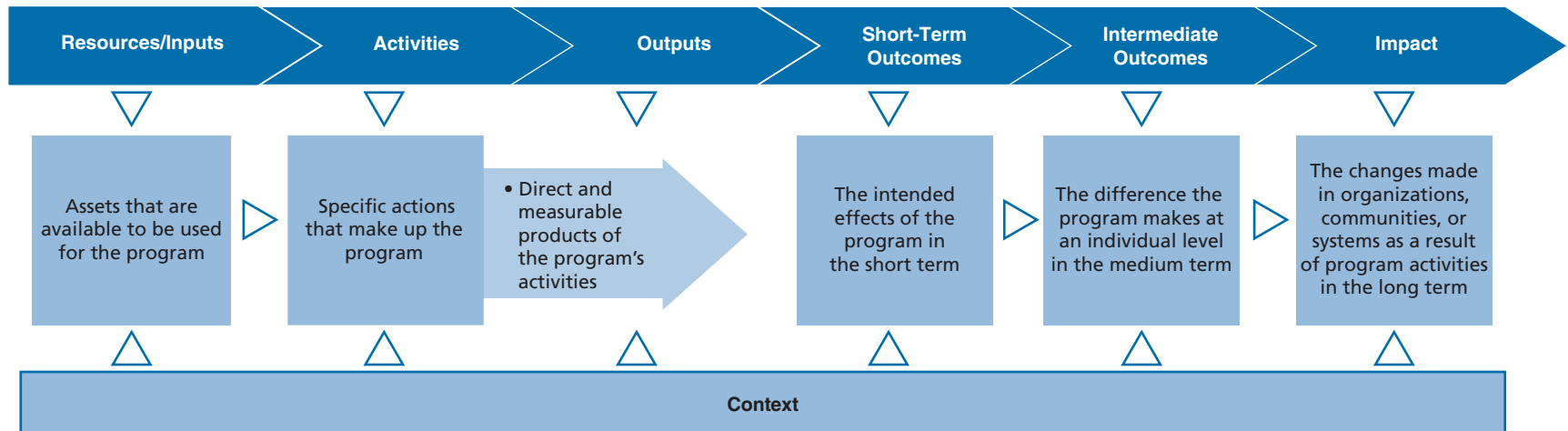
Logic models are a widely used way to visually present the relationships between a PSC's core components and summarize the way that a PSC works. This includes the following elements:

- The **resources/inputs** are the assets necessary for the PSC to occur, including the various human, organizational, financial, and physical resources needed to operate the PSC, support the planned activities, and make implementation possible.
- The **activities** are the specific actions, strategies, and approaches that are implemented to achieve a PSC's goals and bring about intended changes.
- The **outputs** are the direct and tangible products of a PSC's activities.
- The **outcomes** are the intended short-term and intermediate effects of the PSC at the individual level.
- The **impact** refers to the intended aggregate long-term changes in organizations, communities, or systems that result from the PSC.
- The **context** represents the external factors or conditions that influence implementation and the ability to achieve the intended outcomes.

A basic outline of a logic model appears in Figure 3.1. The relationships between the core components are modeled using boxes and arrows. A logic model can help managers justify and build support for a PSC and guide the development of an evaluation plan. It can also help evaluators identify the outcomes that should be the focus of the PSC, or identify aspects that need additional support. For the time frame for outcomes and impact, we use the first 12 months for short-term outcomes, one to three years for intermediate outcomes, and four to six years for impact. However, if your program is only three months long, your short-term outcomes may occur in the one- to three-month time frame, and your intermediate outcomes may occur in the six-month to one-year time frame. These time frames should be adjusted to fit the specific PSC you are evaluating.

In this chapter, we provide guidance on describing the goals of your PSC, identifying and describing the core components of your PSC, and organizing these elements into a logic model to clearly visualize the relationships between components. We also provide suggestions for

**Figure 3.1**  
**Logic Model Overview**



RAND TL264-3.1

additional resources and materials that offer further guidance and support in developing logic models. As noted before, some of the terminology used in this chapter may be different from the terminology that you typically use to describe these concepts, but we will use the terms introduced here consistently throughout this User's Guide.

## Identifying the Goals of Your Program, Service, or Capability

In general, logic models display the connection between your PSC's goals and the core components named earlier. In this way, it can serve as a useful tool to program evaluators. Before mapping out the core components of your logic model, you should make sure that you have clearly defined the goals of your PSC. When identifying your PSC goals, you should follow the SMART principles (Lesesne et al., 2011). That is, they should be Specific, Measurable, Achievable, Relevant, and Time-bound. Here is more detail about what this means in the context of identifying program goals (Acosta et al., 2013):

- **Specific:** Your PSC goal(s) should describe exactly what is expected to change, and for whom.
- **Measurable:** There must be a way to determine whether a goal has been met, the extent of progress that has occurred toward this goal, or change on the outcomes expected as a result of your PSC.
- **Achievable:** The goals of your PSC must be feasible for the target population (e.g., based on prior empirical expectations for change) and with the available resources.
- **Relevant:** The goals should be related to the program purpose and mission.
- **Time-bound:** The time frame in which the change is expected to occur should be specified.

We can examine two examples of PSC goals to demonstrate these principles. Imagine a program whose mission is to prepare Marines for the transition to civilian life. Table 3.1 compares two sample goals for this hypothetical program to determine whether they are consistent with the SMART principles.

With these principles in mind, you can begin to identify the goals of your PSC using Worksheet 3.1. Begin with the overall mission or purpose of your PSC. This will help you as you begin to consider the more specific goals that your PSC is designed to achieve and will also help you to determine whether your goals meet the SMART principles (e.g., by making sure that the goals you identify are relevant to the mission or purpose of your PSC). Then, complete the boxes next to each of the SMART principles. First, you will identify the specific changes or results you expect to occur. Next, consider whether you have identified measurable goals; for instance, how will you know whether the goal is being achieved? Third, consider whether the anticipated goals are achievable for your population and with the resources you have available, and whether they are relevant to the larger mission or purpose of your PSC. Finally, what is the time frame for these goals, and are they anchored to specific milestones (e.g., an overseas move, transitioning from the Marine Corps)? There are questions listed below each SMART principle in this worksheet to assist in this process. Finally, based on your responses on this worksheet, identify your PSC goals.

**Table 3.1**  
**Determining Whether Goals Meet the SMART Principles**

|            | Sample Goal #1: Aid in the development of life skills                                                                                                                                                                                     | Sample Goal #2: Increase Marine knowledge of available resources for finding a job after leaving the USMC                                                                                                                                                                  |
|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Specific   | <input checked="" type="checkbox"/><br>This goal could be made more specific by identifying a specific type or category of life skills, and by identifying a target population.                                                           | <input checked="" type="checkbox"/><br>This goal is specific to one aspect of the transition from the USMC: finding a job.                                                                                                                                                 |
| Measurable | <input checked="" type="checkbox"/><br>Without knowing what aspects of life skills this program aims to develop, it would be challenging to effectively measure this goal.                                                                | <input checked="" type="checkbox"/><br>This goal could be measured by determining what resources Marines are aware of and actually use.                                                                                                                                    |
| Achievable | <input checked="" type="checkbox"/><br>It is realistic to expect that a program within the MF Division could help Marines to develop life skills. However, meeting one of the SMART principles is not enough to have a well-defined goal. | <input checked="" type="checkbox"/><br>This does appear to be a realistic goal for a program, as the scope is reasonable. This might be less achievable if, for instance, the goal was to make sure that Marines know of <i>all</i> available resources for finding a job. |
| Relevant   | <input checked="" type="checkbox"/><br>This goal might be relevant to the broader program mission of facilitating the transition to civilian life, but without knowing the specific life skills, this cannot be determined conclusively.  | <input checked="" type="checkbox"/><br>This goal is directly relevant to the program's mission of facilitating the transition to civilian life.                                                                                                                            |
| Time-bound | <input checked="" type="checkbox"/><br>There is no specific time frame identified. Are these skills that will be developed before leaving the Marines that will be helpful down the road? Or skills for navigating the transition?        | <input checked="" type="checkbox"/><br>This goal is time-bound—it is tied to the time frame of the transition, and the acquisition of knowledge could be considered a short-term goal of this program.                                                                     |

### Worksheet 3.1 Identifying SMART Goals

|                                                                                                                                                                                                                                     |  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| What is the mission or purpose of your PSC?                                                                                                                                                                                         |  |
| <hr/>                                                                                                                                                                                                                               |  |
| <hr/>                                                                                                                                                                                                                               |  |
| Specific <ul style="list-style-type: none"> <li>What is expected to change?</li> <li>How will this be accomplished?</li> <li>Who will experience this change?</li> </ul>                                                            |  |
| Measurable <ul style="list-style-type: none"> <li>How will you measure progress toward this outcome?</li> <li>How will you know whether this goal is being achieved?</li> </ul>                                                     |  |
| Achievable <ul style="list-style-type: none"> <li>Are these changes feasible for the population?</li> <li>Does your program have the resources available to support these changes?</li> </ul>                                       |  |
| Relevant <ul style="list-style-type: none"> <li>Are the changes you identified related to the mission or purpose of your PSC?</li> <li>Are they related to the activities your PSC offers?</li> </ul>                               |  |
| Time-bound <ul style="list-style-type: none"> <li>What is the time frame in which you expect to observe effects?</li> <li>Are your PSC's activities tied to specific milestones or events for Marines or family members?</li> </ul> |  |
| Based on your responses to the questions above, identify your PSC goal(s):                                                                                                                                                          |  |
| 1. <hr/>                                                                                                                                                                                                                            |  |
| 2. <hr/>                                                                                                                                                                                                                            |  |
| 3. <hr/>                                                                                                                                                                                                                            |  |
| 4. <hr/>                                                                                                                                                                                                                            |  |

## Identifying Core Components

In this section, we will briefly introduce each component in a logic model and then walk you through a process to help you identify and describe the core components of your PSC.

For each core component, we have provided some questions to consider as you gather information and identify key aspects of how things operate. These are suggested questions rather than an exhaustive list and will differ by the type of PSC.

Each section below has a worksheet to organize the information gathered. The worksheets are designed to help you systematically collect the necessary information to build your logic model. We also provide examples of how to complete portions of the worksheet for the example program (L.I.N.K.S.), service (FMEAP), and capability (Outdoor Recreation Equipment Checkout and Equipment Rental) described in the preceding chapter.

As we describe the logic model, we will work from left to right on the logic model depicted in Figure 3.1—that is, starting with resources/inputs and moving to outcomes and impact. However, in some situations, it may be helpful to complete the logic model in the reverse order. For example, if you are just beginning to develop a PSC, you may know what outcomes and impact you want to achieve. From there, you might then identify what your outputs should

be, and then identify what activities would be needed to produce those outputs, and ultimately identify the resources and inputs necessary to make that possible. In this case, it may be helpful to read through the text of this chapter in the order that it is presented as a comprehensive overview, but then complete the worksheets in the reverse order to develop your logic model.

## Resources/Inputs

The logic model starts with the resources needed to implement the PSC. Resources are the assets or *inputs* needed and available to support the operation of the program. The performance of activities and the fulfillment of goals depend on these resources. A successful PSC needs adequate resources to fulfill needs, as well as prudent allocation of these resources to avoid wasting time and money.

The questions below will help you identify some of the most common resources that your PSC may use, including staff, facilities, funding, and other resources.

### Worksheet 3.2 Resources/Inputs

| Type       | Description |
|------------|-------------|
| Staff      |             |
| Facilities |             |
| Funding    |             |
| Other      |             |

**Example Resources/Inputs for L.I.N.K.S. Program**

| Type  | Description                                                                                                                                                                                                                                                                                                   |
|-------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Staff | <ul style="list-style-type: none"> <li>At least one full or part-time L.I.N.K.S. trainer to coordinate all training, supervise an administrative assistant, and recruit, train, and supervise all volunteers.</li> <li>Mentor volunteers who have attended a session and received mentor training.</li> </ul> |

**Example Resources/Inputs for FMEAP Service**

| Type    | Description                                                                                                                                                                                                                                                                                                                   |
|---------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Funding | <ul style="list-style-type: none"> <li>Mix of appropriated and converted nonappropriated funds to provide adequate funding for personnel, training, maintenance, telephone lines, internet access, scanners and copiers, supplies for automated equipment, and other support necessary for the operation of FMEAP.</li> </ul> |

**Example Resources/Inputs for Outdoor Recreation Equipment Checkout and Equipment Rental Capability**

| Type       | Description                                                                                                    |
|------------|----------------------------------------------------------------------------------------------------------------|
| Facilities | <ul style="list-style-type: none"> <li>Accessible space to store and maintain the rental equipment.</li> </ul> |

- Staff
  - How many and what type of staff are needed to implement and support the PSC (e.g., trainer, manager)?
  - What staff expertise is required?
  - What training or credentials do staff need?
- Facilities
  - What facilities are used or needed for activities (e.g., classroom, office)?
  - What technology will be utilized in the PSC (e.g., computers for webinars)?
- Funding
  - What is the actual and needed funding level for this PSC?
  - Where does the funding come from?
- Other resources
  - What other resources does the PSC use/need for its operational purposes (e.g., purchased training materials, testing equipment, technologies, external contractors)?

After considering these questions, complete Worksheet 3.2 to identify the resources/inputs for your PSC. Be specific, as your logic model will ultimately be a tool for planning your evaluation. The more detail you provide when developing your logic model, the easier it will be to design your evaluation (Chapter Four) and identify evaluation measures (Chapter Five).

Next, we provide examples of how the worksheet descriptions can be used to develop a well-specified description of resources for the logic model.

Worksheet 3.3  
Activities

| Type | Frequency/Duration | Staff Needed | Intended Recipient |
|------|--------------------|--------------|--------------------|
|      |                    |              |                    |
|      |                    |              |                    |
|      |                    |              |                    |
|      |                    |              |                    |

**Example Activity for L.I.N.K.S. Program**

| Type                                     | Frequency/Duration             | Staff Needed                  | Intended Recipient |
|------------------------------------------|--------------------------------|-------------------------------|--------------------|
| L.I.N.K.S. training sessions for spouses | Two 4.5-hour training sessions | L.I.N.K.S. Trainer and Mentor | Marine spouses     |

**Example Activity for FMEAP Service**

| Type                                                                                    | Frequency/Duration | Staff Needed                            | Intended Recipient                            |
|-----------------------------------------------------------------------------------------|--------------------|-----------------------------------------|-----------------------------------------------|
| Employment resources, such as job search sites and resource lists, referrals, job fairs | Ongoing, as needed | Installation FMEAP Managers/Specialists | Military spouses and dependent family members |

**Example Activity for Outdoor Recreation Equipment Checkout and Equipment Rental Capability**

| Type                                                                                                  | Frequency/Duration    | Staff Needed                                                                                                            | Intended Recipient                            |
|-------------------------------------------------------------------------------------------------------|-----------------------|-------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------|
| Provide equipment checkout and rental for self-directed activities, instructional classes, and events | Ongoing, as requested | Program Manager and Recreation Assistant, Maintenance Mechanic, and Recreation Specialist Programmer, depending on need | Active-duty Marines, dependent family members |

**Activities**

Activities are the specific actions or strategies that make up the PSC and are employed to reach goals. This includes the tools, services, and products that the PSC provides to the target population. Some questions to consider when identifying and describing your PSC's activities:

- What activities (e.g., training sessions, events, meetings, counseling sessions, or other services and supports) are part of this PSC?
- What is the frequency and duration of each activity (e.g., weekly for one hour, six-hour annual training)?
- Who delivers or supports the implementation of the activity?
- Who is the recipient of each activity (e.g., active-duty Marine, spouse, family)?

After considering these questions, complete Worksheet 3.3 to identify the activities for your PSC.

**Worksheet 3.4**  
**Target Population**

| Description (e.g., age, gender, position, area, risk factor) | Eligibility Criteria | Identification and Recruitment |
|--------------------------------------------------------------|----------------------|--------------------------------|
|                                                              |                      |                                |
|                                                              |                      |                                |
|                                                              |                      |                                |
|                                                              |                      |                                |

Next, we provide examples of how the worksheet descriptions can be used to develop a well-specified list of activities for the logic model.

**Example Target Population for L.I.N.K.S. Program**

| Description (e.g., age, gender, position, area, risk factor)                                                                                                                    | Eligibility Criteria                                                               | Identification and Recruitment                                                                                          |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|
| All Marines, spouses, school-aged children, teens, and parents/extended family; other service members and spouses serving with Marines; and USMC civilian employees and spouses | Program primarily targets new Marines and families, but the program is open to all | Self-referral. Information about programs is available on installation Marine Corps Community Services (MCCS) websites. |

**Example Target Population for FMEAP Service**

| Description (e.g., age, gender, position, area, risk factor) | Eligibility Criteria                                                          | Identification and Recruitment |
|--------------------------------------------------------------|-------------------------------------------------------------------------------|--------------------------------|
| Military spouses                                             | Married to an active-duty family member (reservists, retirees, and civilians) | Self-referral                  |

**Example Activity for Outdoor Recreation Equipment Checkout and Equipment Rental Capability**

| Description (e.g., age, gender, position, area, risk factor) | Eligibility Criteria | Identification and Recruitment |
|--------------------------------------------------------------|----------------------|--------------------------------|
| Active-duty Marines; military family and dependents          | None                 | Self-referral                  |

**Target Population**

The target population is the group of individuals that the PSC intends to serve with its activities. This may consist of the entire population of Marines and their families or a subgroup of this population defined by the eligibility criteria. The questions below will help you be as specific as possible when describing the PSC's target population.

- Whom does this program intend to serve each year?
  - What is the age range (youth, adults), gender, or occupation of target participants?
- What are the specific eligibility criteria for participating in the PSC?
- How are potential participants in the PSC identified and recruited?
  - Who identifies and refers potential participants?
  - What type of outreach is conducted to recruit potential participants?

After considering these questions, complete Worksheet 3.4 to list the target population(s) for your PSC.

**Worksheet 3.5****Outputs**

| Activity | Outputs<br>(e.g., amount or volume of activity, quality, satisfaction) |
|----------|------------------------------------------------------------------------|
|          |                                                                        |
|          |                                                                        |
|          |                                                                        |
|          |                                                                        |
|          |                                                                        |

Next, we provide examples of how the worksheet descriptions can be used to develop a well-specified description of the target population for the logic model.

**Example Outputs for L.I.N.K.S. Program**

| Activity                                 | Outputs<br>(e.g., amount or volume of activity, quality, satisfaction)                                                                                                                                                                                                                                                                                  |
|------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| L.I.N.K.S. training sessions for spouses | <ul style="list-style-type: none"> <li>• Number of sessions offered per year, by type of session</li> <li>• Number of families who attend any type of L.I.N.K.S. session</li> <li>• Proportion of interested families attending a session</li> <li>• Average attendance, by type of session</li> <li>• Participant satisfaction with program</li> </ul> |

**Example Outputs for FMEAP Service**

| Activity                                                                        | Outputs<br>(e.g., amount or volume of activity, quality, satisfaction)                                                                                                                                                                                                                                                                        |
|---------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Employment resources, job search sites and resource lists, referrals, job fairs | <ul style="list-style-type: none"> <li>• Number of spouses and dependents attending career training or education sessions</li> <li>• Number of job search sites and local job opportunities</li> <li>• Number of referral service appointments</li> <li>• Number of job fairs hosted and number of companies participating in each</li> </ul> |

**Example Outputs for Outdoor Recreation Equipment Checkout and Equipment Rental Capability**

| Activity                                                                                              | Outputs<br>(e.g., amount or volume of activity, quality, satisfaction)                                                                                                                                                                                                                                                                                                                                                                                    |
|-------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Provide equipment checkout and rental for self-directed activities, instructional classes, and events | <ul style="list-style-type: none"> <li>• Number of weekly hours per activity type</li> <li>• Number of available required equipment pieces</li> <li>• Number of checkouts by individuals, by type, cost, and value of equipment</li> <li>• Number of skills and safety lessons provided</li> <li>• Number of individuals receiving skills and safety lessons</li> <li>• Proportion of individuals receiving required skills and safety lessons</li> </ul> |

**Outputs**

Outputs are the amount, quality, and/or volume of each activity within a PSC. These can include the number of individuals trained, number of individuals tested, the quantities of materials developed or distributed, or participant satisfaction.

The following questions can help identify the direct products of each PSC activity that you may be interested in measuring.

- How many participants from the target population participated during (insert time frame)?
  - What percentage of the eligible population participated in the PSC during (insert time frame)?
  - What percentage of those identified and referred to the program, service, or capability participated during (insert time frame)?
- How much of the activity was provided during (insert time frame)?
  - What was the average amount of the activity each participant received?
- How were the activities received by participants? Were they satisfied with the scope and quality of the program, service, or capability?

After considering these and other related questions that may come to mind, complete Worksheet 3.5 to list the outputs from each activity for your PSC.

Next, we provide examples of well-specified descriptions of the outputs associated with the activities offered within the context of our three PSC examples. Note that these lists are not exhaustive; they intend to provide examples of the outputs that might be expected from each PSC.

## Outcomes

Outcomes are changes in the target population expected as a result of engaging in the activities. Outcomes often include individual-level changes in knowledge, attitudes, skills, or behavior as a result of the PSC. Effective outcomes follow the SMART principles that were introduced

### Worksheet 3.6 Outcomes

| Activity | Outcome<br>(What is expected to change?) | What target population is expected to change? | In what time frame will the change occur? |
|----------|------------------------------------------|-----------------------------------------------|-------------------------------------------|
|          |                                          |                                               |                                           |
|          |                                          |                                               |                                           |
|          |                                          |                                               |                                           |
|          |                                          |                                               |                                           |

earlier—they should be specific, measurable, achievable, relevant, and time-bound (Lesesne et al., 2011).

These outcomes represent the intended effects of the PSC in the short and intermediate term. Identifying outcomes is important because, ultimately, a goal of your evaluation will be to understand and demonstrate how the PSC has resulted in changes in the individual-level outcomes of interest.

#### Example Outcomes for L.I.N.K.S. Program

| Activity                                                        | Outcome<br>(What is expected to change?)                                                | What target population is expected to change? | In what time frame will the change occur? |
|-----------------------------------------------------------------|-----------------------------------------------------------------------------------------|-----------------------------------------------|-------------------------------------------|
| L.I.N.K.S. training sessions for spouses (short-term outcomes)  | Increased basic military knowledge (e.g., deployment, rank structure, chain of command) | Spouses                                       | By conclusion of workshop                 |
|                                                                 | Increased understanding of USMC mission, organization, and culture                      | Spouses                                       | By conclusion of workshop                 |
| L.I.N.K.S. training sessions for spouses (intermediate outcome) | Increased social support from other military families (e.g., social connection)         | Spouses                                       | Six months to one year after workshop     |

#### Example Outcomes for FMEAP Service

| Activity                                                                                               | Outcome<br>(What is expected to change?)                                                                                                                     | What target population is expected to change? | In what time frame will the change occur?             |
|--------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------|-------------------------------------------------------|
| Employment resources, job search sites and resource lists, referrals, job fairs (short-term outcome)   | Follow-up on referrals/resources provided by program staff                                                                                                   | Family members                                | Within three months of receiving referral             |
|                                                                                                        | Increased knowledge or skills regarding career-related topics (e.g., resume content, conducting a job search, interview skills)                              | Family members                                | At conclusion of training session or coaching session |
|                                                                                                        | Completed resume                                                                                                                                             | Family members                                | At conclusion of training session or coaching session |
|                                                                                                        | Prepared for job interviews                                                                                                                                  | Family members                                | At conclusion of training session or coaching session |
| Employment resources, job search sites and resource lists, referrals, job fairs (intermediate outcome) | Measurable progress toward career goals (specific outcomes will depend on career goal but may include submitting job applications, obtaining job interviews) | Family members                                | Six months to one year after services received        |

**Example Outcomes for Outdoor Recreation Equipment Checkout and Equipment Rental Capability**

| <b>Activity</b>                                                                                                              | <b>Outcome<br/>(What is expected to change?)</b>       | <b>What target population is expected to change?</b> | <b>In what time frame will the change occur?</b>                                                            |
|------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------|------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|
| Provide equipment checkout and rental for self-directed activities, instructional classes, and events (short-term outcome)   | Access to affordable/healthy leisure activities        | Marines and family members                           | Immediate—upon arrival at Equipment Rental                                                                  |
|                                                                                                                              | Skills for outdoor recreation activities and equipment | Marines and family members                           | At conclusion of skills and safety lesson or immediately following equipment rental                         |
| Provide equipment checkout and rental for self-directed activities, instructional classes, and events (intermediate outcome) | Increased perceptions of community connectedness       | Marines and family members                           | Six months after equipment rental. May observe to greater degree among those who use capability more often. |

The following questions can help identify the intended effects of your PSC in the short and intermediate terms. As noted earlier, in this toolkit we have defined short-term outcomes as those expected in the first 12 months and intermediate outcomes as those expected in one to three years. Since implementation periods differ widely, these time frames should be adjusted to fit the specific PSC you are evaluating.

- Short-term outcomes:
  - What changes do you expect participants to experience in the short term (e.g., new knowledge, skills, abilities, certifications, or connections)?
  - What need is this PSC designed to meet in the short term?
  - What target population is expected to change?
  - In what time frame do you expect the change to occur?
- Intermediate outcomes:
  - What changes do you expect participants to experience in the intermediate term (e.g., after financial skills training sessions, Marines will demonstrate financial well-being through reduced debts and increased savings)?
  - What issues are this PSC designed to target in the intermediate term?
  - What target population is expected to change?
  - In what time frame do you expect the change to occur?

After considering these and other related questions, complete Worksheet 3.6 to identify the outcomes for your PSC.

Next, we provide examples of how the worksheet descriptions can be used to develop a well-specified list of outcomes for the logic model.

**Worksheet 3.7**  
**Impact**

| Short- or intermediate-term outcome | Expected impact at the aggregate level (e.g., unit, total force, installation) |
|-------------------------------------|--------------------------------------------------------------------------------|
|                                     |                                                                                |
|                                     |                                                                                |
|                                     |                                                                                |
|                                     |                                                                                |
|                                     |                                                                                |

**Impact**

The logic model also lays out the intended impact of your PSC over time. These impacts occur

**Example Impact for L.I.N.K.S. Program**

| Short- or Intermediate-Term Outcome       | Expected impact at the aggregate level (e.g., unit, total force, installation)       |
|-------------------------------------------|--------------------------------------------------------------------------------------|
| Social support from other Marine families | Increased Marine and family resilience<br>Improved Marine and family quality of life |

**Example Impact for FMEAP Service**

| Short- or Intermediate-Term Outcome                                                                                             | Expected impact at the aggregate level (e.g., unit, total force, installation) |
|---------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| Increased knowledge or skills regarding career-related topics (e.g., resume content, conducting a job search, interview skills) | Increased Marine and family retention                                          |

**Example Impact for Outdoor Recreation Equipment Checkout and Equipment Rental Capability**

| Short- or Intermediate-Term Outcome             | Expected impact at the aggregate level (e.g., unit, total force, installation) |
|-------------------------------------------------|--------------------------------------------------------------------------------|
| Access to affordable/healthy leisure activities | Increased Marine and family readiness                                          |

at the aggregate level—that is, at the unit, installation, or total-force level.<sup>1</sup> Across the MF portfolio, PSCs vary widely in their focus and activities. However, they are largely designed to increase readiness, improve resilience, increase retention in the service, and improve quality of life for service members and their families. In addition, certain PSCs have intended impacts in other domains, such as supporting morale, welfare, and recreation.

When identifying the expected impact of your program, it can be helpful to envision the short-term and intermediate outcomes that you identified previously and then consider what those outcomes may result in when aggregated at the unit, installation, or total-force level over the long term. The following questions can also help identify the intended impact of your PSC in the long term.

- What are the intended long-term outcomes or changes?
  - What is expected to change at the battalion, company, or division level when individual Marines participate in the PSC (e.g., increased morale)?
  - What is expected to change at a USMC installation?
  - How does the USMC, as an organization, benefit from operating this PSC?

After considering these and other related questions, complete Worksheet 3.7 to identify the impact areas for your PSC.

Next, we provide examples of how the worksheet descriptions can be used to develop a well-specified list of impact areas for the logic model.

<sup>1</sup> Outcomes may be measured at the individual, family, or unit level. When aggregated, whether across the total force (e.g., active-duty Marines) or installations, these outcomes are considered impacts.

**Worksheet 3.8**  
**Context**

| Contextual Factor | Area of Impact |
|-------------------|----------------|
|                   |                |
|                   |                |
|                   |                |
|                   |                |
|                   |                |

**Context**

Logic models can also contain a section on local and external context. These are the political,

**Example Context for L.I.N.K.S. Program**

| Contextual Factor                                                      | Area of Impact |
|------------------------------------------------------------------------|----------------|
| Different installations offer various types of L.I.N.K.S. programming. | Outputs        |

**Example Context for FMEAP Service**

| Contextual Factor                                                                                                                                                                | Area of Impact |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|
| FMEAP coordinates with community employers, MCCS, and Family Readiness programs to provide referrals, resources, and possible volunteer opportunities to spouses and dependents. | Outputs        |

**Example Context for Outdoor Recreation Equipment Checkout and Equipment Rental Capability**

| Contextual Factor                                                                                                                                                                                                                                                                                                                                                                                                                 | Area of Impact                 |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|
| Outdoor Recreation Equipment Checkout and Equipment Rental supports Directed Outdoor Recreation (an MF program). This capability also provides unit support and unit-sponsored recreational events defined as holidays or special parties held in conjunction with official events (e.g., award ceremonies, changes of command). Unit support is provided as a Mission Sustaining morale, welfare, and recreation (MWR) function. | Resources/Inputs<br>Activities |
| Available equipment varies across installations.                                                                                                                                                                                                                                                                                                                                                                                  | Resources/Inputs               |

social, cultural, or economic situations within the USMC or at specific installations that can facilitate or hinder the utilization of MF PSCs. Since the PSC may work differently in different settings, the logic model can identify contextual factors that may influence both implementation and the subsequent outcomes and explain why those outcomes differ across settings.

The questions below will help you identify the contextual factors that make up this core component.

- Has there been a change (increase or decrease) in support for this PSC from USMC leadership?
- Are there competing or supporting initiatives within or external to the PSC that may decrease or increase expected participation?
- What are the social norms about participating in this PSC that may affect participation and effectiveness?
- Are there other barriers or supports that shape the PSC's implementation or functioning?

Certain aspects of the military experience may also be an important part of the context. For example, do deployments affect an individual's ability to fully engage with your PSC?

After considering these and other related questions, complete Worksheet 3.8 to identify the context for your PSC. Although these contextual factors will not necessarily be explicitly measured as part of your evaluation, they will help you to interpret your evaluation findings.

Next, we provide examples of how the worksheet descriptions can be used to develop a well-specified description of the context for the logic model.

## Building the Logic Model

Logic models facilitate the visualization of relationships between the core components of a PSC. Reading from left to right, these models depict the logical flow of inputs (resources) to actions (activities) to outputs to outcomes and impact. An “if . . . then” relation connects adjacent components: If the resources are provided, then the activities are conducted; if the activities are conducted, then the targeted population can participate; if the targeted population participates, then the outcomes can occur.

The organization of a logic model has the following advantages:

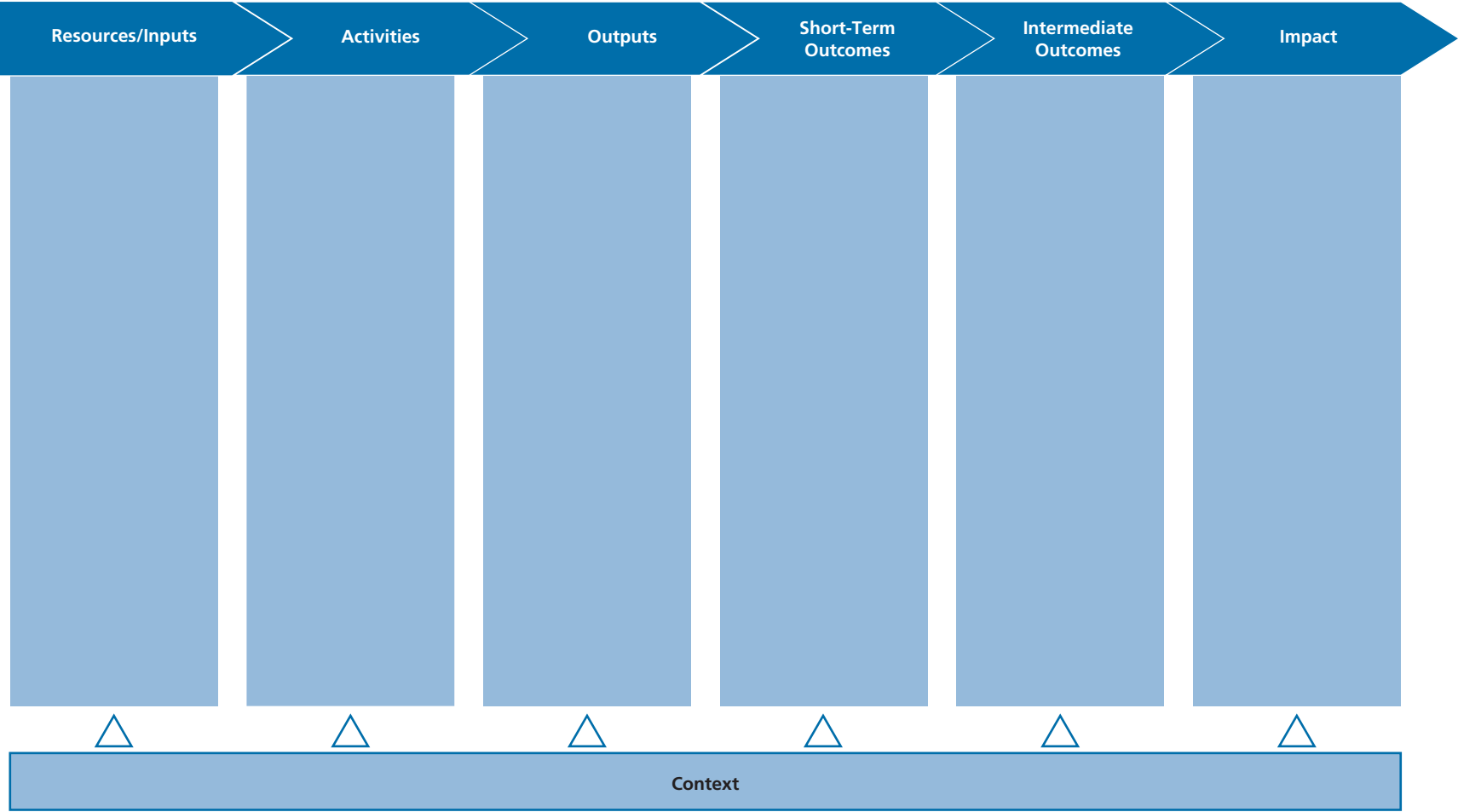
- It simplifies and summarizes the core components.
- It depicts the connection between concrete resources/activities and abstract goals.
- It allows for assessment of how well the plan aligns with the actual implementation.

The template following (Worksheet 3.9) provides a structure for organizing the logic model information and drawing the connections between each of the key components. Now that you have identified and specified the core components of the PSC, you can use the responses from Worksheets 3.2 to 3.8 to construct a logic model using Worksheet 3.9. The well-specified descriptions in Worksheets 3.2 to 3.8 can be transferred directly into the corresponding section of the logic model in Worksheet 3.9.

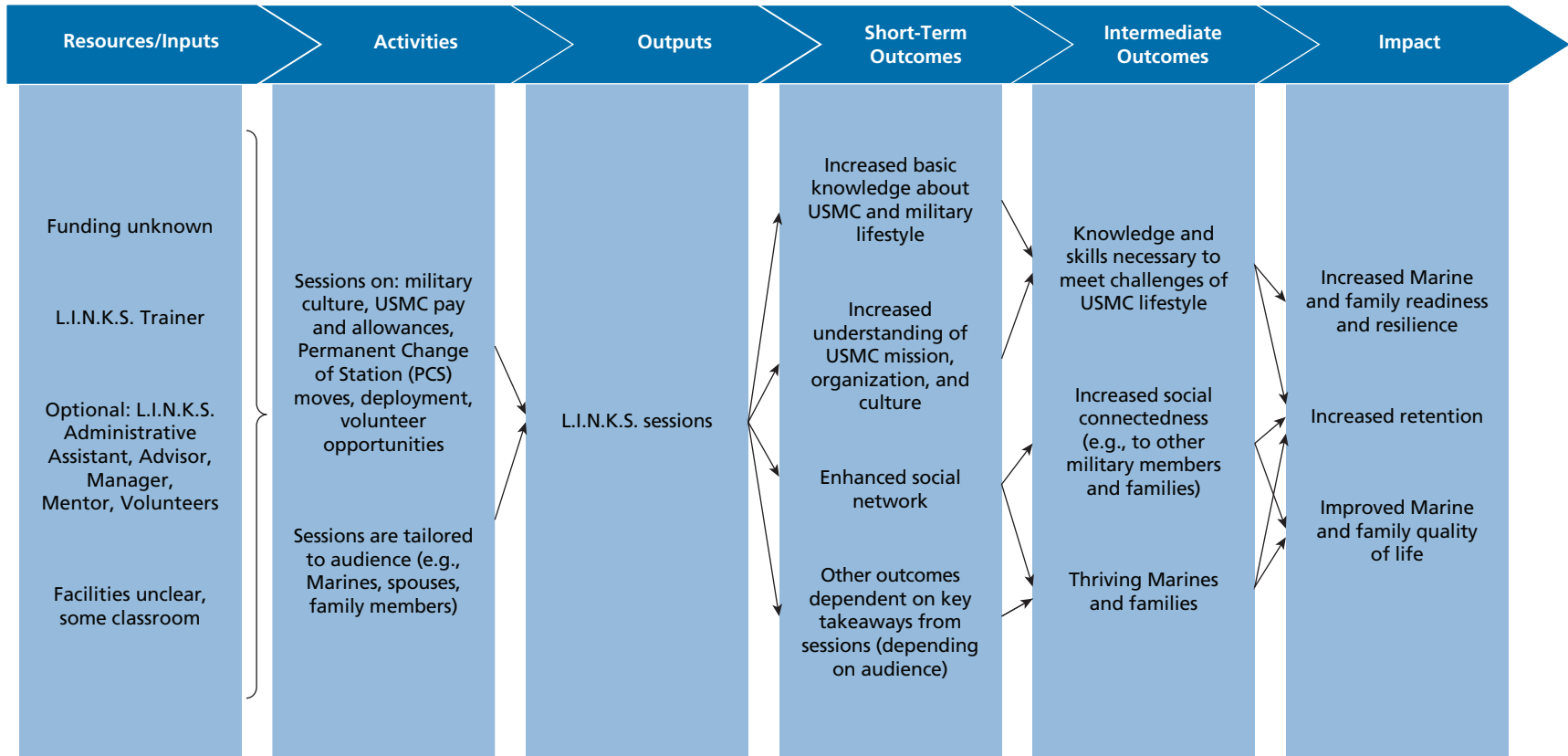
Using this basic structure, you can add arrows and specify boxes to identify the interrelationship between elements of the major components. The connections should follow resources through to impact. Identifying these interrelationships can be a helpful step when developing your logic model. Not only does it help you to more clearly consider the links between your outputs and your outcomes—both actual and desired—but it helps you to identify any potential disconnections. For example, you might realize that you have identified an intended outcome of your PSC, but that none of the outputs are connected to that outcome. If the program activities do not support the outcomes, then you may not observe the intended effects. Therefore, identifying these disconnections can be an important step toward making sure that your PSC can achieve its goals. Next, we provide completed logic models for the example program (Figure 3.2, L.I.N.K.S.), service (Figure 3.3, FMEAP), and capability (Figure 3.4, Outdoor Recreation Equipment Checkout and Equipment Rental) used throughout this chapter.

Logic models can be dynamic. As you develop the logic model assumptions or discuss how the PSC functions in a specific context, more information can be added, or connections between components can be altered. Additionally, the program logic may change over time when resources, activities, or contextual factors also change. Dating a logic model will help you track these changes throughout the life of your PSC.

Worksheet 3.9  
The Logic Model

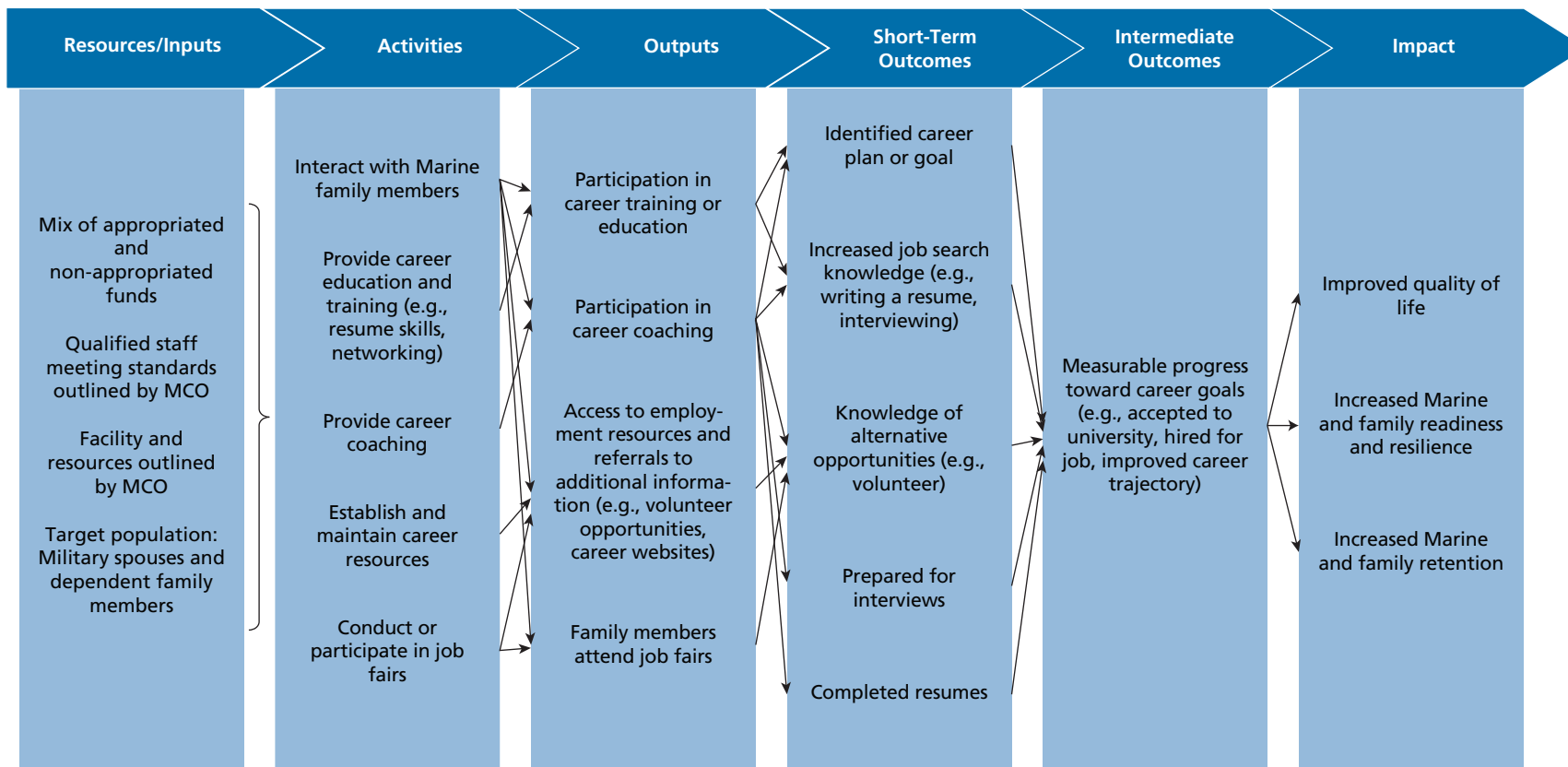


## Logic Model for L.I.N.K.S. Program



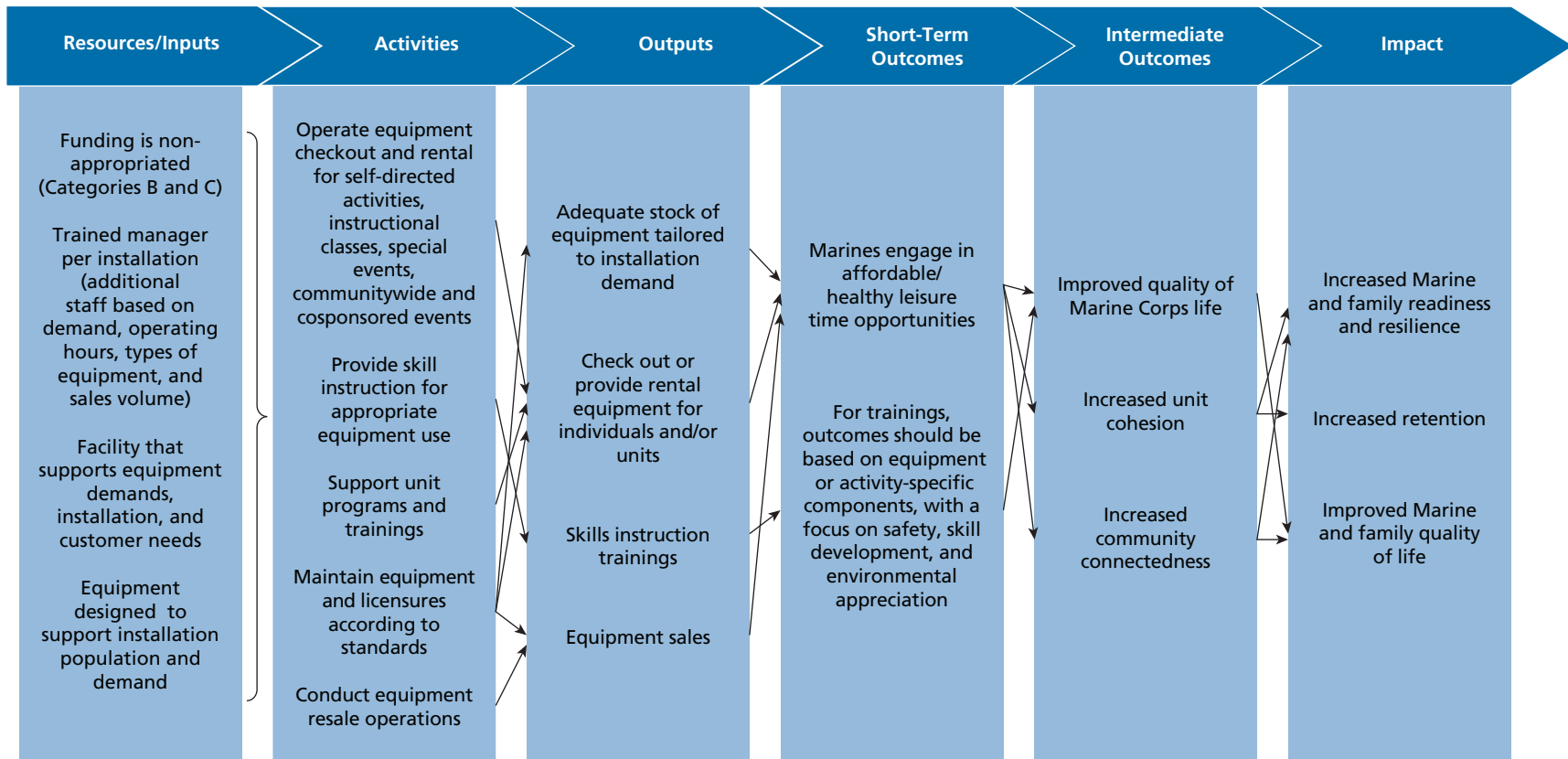
RAND TL264-3.2

## Logic Model for FMEAP Service



RAND TL264-3.3

## Logic Model for Outdoor Recreation Equipment Checkout and Equipment Rental Capability



RAND TL264-3.4

## Using the Logic Model for Your Evaluation

Ultimately, your completed logic model will form the basis for your evaluation. In the next chapter, we will discuss two types of evaluations: process evaluations and outcome evaluations. Generally speaking, process evaluations focus on the first three components of your logic model, including resources/inputs, activities, and outputs, while outcome evaluations focus on the short-term and intermediate individual-level outcomes of your program, as well as longer-term outcomes or impact at the aggregate level. As part of the evaluation planning process, you will learn to identify measures of performance, which quantify the implementation and utilization of your PSC, and measures of effectiveness, which help you to determine whether your program had the intended effects. Developing a detailed logic model is an important first step, though, because it ensures that you have specifically defined the resources/inputs, activities, outputs, and intended effects of your PSC. The next step is to determine how to measure these elements. The more detailed your logic model, the easier this process will be. We will provide more detail about this next step, including worksheets to help you identify your own measures of performance and effectiveness in Chapter Five.

## Additional Guidance

Should your team want more guidance for developing a logic model, understanding the critical components of a model, or reviewing examples of education-specific program logic models, here are a few recommended resources to assist you:

- Acosta, J., R. Ramchand, A. Becker, A. Felton, and A. Kofner, *RAND Suicide Prevention Program Evaluation Toolkit*, Santa Monica, Calif.: RAND Corporation, TL-111-OSD, 2013.
- Lawton, B., P. R. Brandon, L. Cicchinelli, and W. Kekahio, *Logic Models: A Tool for Designing and Monitoring Program Evaluations* (REL 2014–007), Washington, D.C.: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Pacific, 2014.
- W. K. Kellogg Foundation, *Logic Model Development Guide*, Battle Creek, Mich., 2004b. <https://www.wkkf.org/resource-directory/resource/2006/02/wk-kellogg-foundation-logic-model-development-guide>
- Wholey, J. S., H. P. Hatry, and K. E. Newcomer, eds., *Handbook of Practical Program Evaluation* (3rd ed.), San Francisco, Calif.: Jossey-Bass, 2010.

## Summary

You should now be able to identify the resources, activities, and outputs linked to your PSC. You have also had the chance to identify the expected short-term and intermediate outcomes, and how these contribute to the long-term outcomes and intended impact of the PSC. You have identified the local and external contexts that may affect the implementation or the outcomes of your PSC. In the next chapter, you will learn about different evaluation approaches,

and how to use information about your PSC and the available resources, expertise, time available for the evaluation, and other considerations to select the most appropriate approach for your own evaluation.



## Understanding Evaluation and Planning for Your Own Evaluation

---

Evaluations generally can be categorized as process evaluations or outcome evaluations.<sup>1</sup> Process evaluations focus on how a PSC was implemented, and whether it was delivered as intended. Outcome evaluations focus on the outcomes and impacts that a PSC is designed to achieve. Figure 4.1 shows how process and outcome evaluation map onto the core components of the logic model you developed in Chapter Three.

This chapter describes the design, focus, and types of process and outcome evaluations. The goal of this chapter is to help you select a type of evaluation design and develop it expressly for your PSC.<sup>2</sup>

### Human Subjects Considerations

Before you begin to plan for your evaluation, it is critical to understand the DoD and Marine Corps guidance related to the protection of human subjects. For example, there are certain distinctions between “quality improvement” activities and human subjects research. Quality improvement activities are designed to monitor quality and improve services. The purpose is *not* to generate conclusions that can be generalized to other programs or populations; rather, the focus is on the specific program and how its quality can be maintained. By contrast, the

---

<sup>1</sup> We also introduced the concept of impact in the previous chapter but will discuss the impact evaluation as it relates to strategic portfolio management (SPM) in Chapter Seven.

<sup>2</sup> Throughout this chapter, we refer to various resources related to program evaluation. For those who are interested, these additional references about program evaluation may be helpful:

Cook, T. D., and D. T. Campbell, *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Boston, Mass.: Houghton Mifflin, 1979.

Fitzpatrick, J. L., J. R. Sanders, and B. R. Worthen, *Program Evaluation: Alternative Approaches and Practical Guidelines*, 4th ed., Boston, Mass.: Pearson, 2011.

Patton, Q., *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*, 4th ed., Thousand Oaks, Calif.: Sage Publications, 2015.

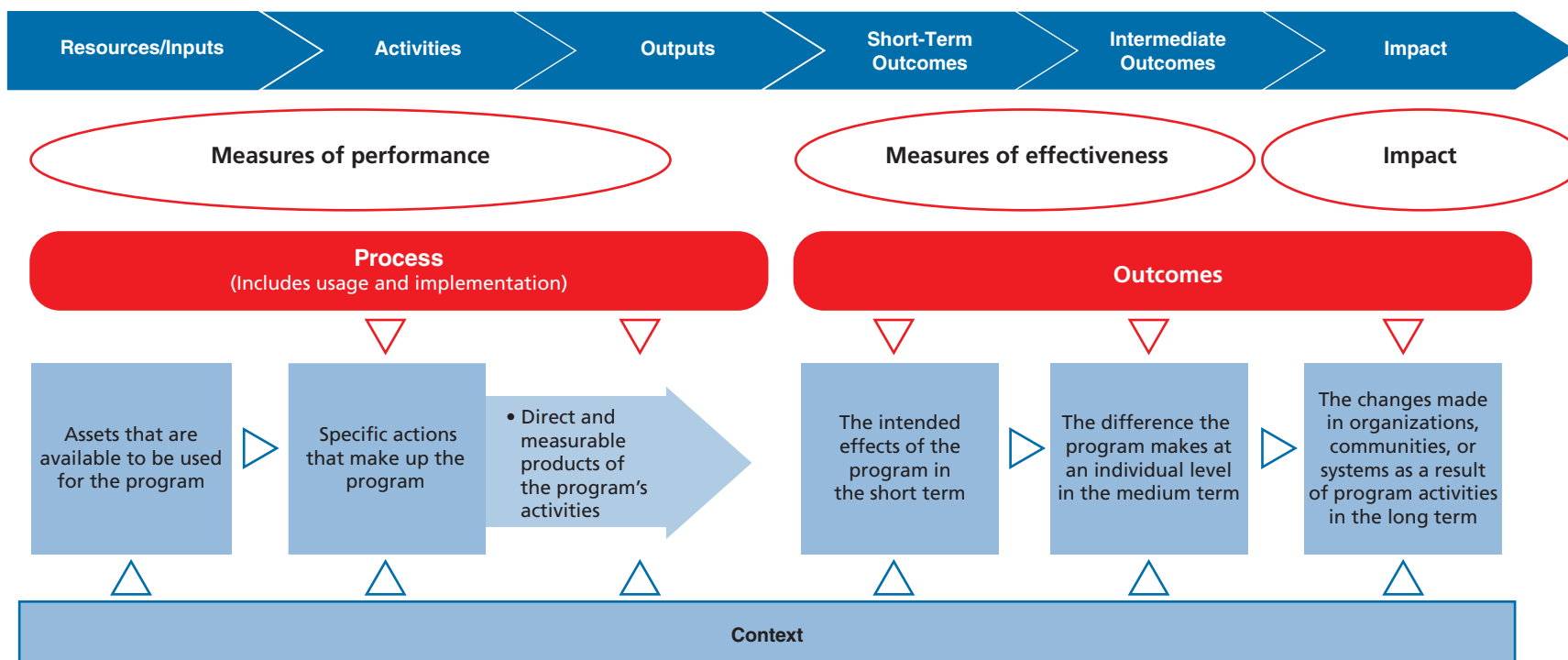
Rossi, P. H., M. W. Lipsey, and H. E. Freeman, *Evaluation: A Systematic Approach*, 7th ed., Thousand Oaks, Calif.: Sage Publications, 2004.

Shadish, W. R., T. D. Cook, and D. T. Campbell, eds., *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, 2nd ed., Boston, Mass.: Houghton Mifflin, 2002.

Shadish, W. R., T. D. Cook, and L. C. Leviton, *Foundations of Program Evaluation: Theories of Practice*, Newbury Park, Calif.: Sage Publications, 1991.

Wholey, J. S., H. P. Hatry, and K. E. Newcomer, eds., *Handbook of Practical Program Evaluation*, 3rd ed., San Francisco, Calif.: Jossey-Bass, 2010.

**Figure 4.1**  
**Framework for Evaluating a Program, Service, or Capability**



RAND TL264-4.1

purpose of human subjects research is to develop generalizable knowledge through a process of systematic investigation (Department of Defense Instruction [DoDI] 3216.02).

Program evaluation activities may fall under the designation of quality improvement as long as “the results of the evaluation are only for the use of Government officials responsible for the operation or oversight of the program being evaluated and are not intended for generalized use beyond such program” (DoDI 3216.02). Whether your evaluation falls under the designation of quality improvement is not a determination that should be made on your own. You should follow the MF-designated Human Subjects Research Protection Program (HRPP) process (see Figure 4.2). All formal mechanisms to solicit feedback from program participants, staff, commands, and other stakeholders (e.g., research, surveys, questionnaires, course evaluations, suggestion box materials) are vetted through the USMC HRPP Applicability Review Process *prior* to use. HRPP provides study determinations concerning whether research, surveys, questionnaires, course evaluations, and suggestion box materials are in accordance with ethical principles, laws, regulations, and policies to ensure respect for persons, beneficence, and justice for human subjects. Program evaluations are considered aspects of research and *must* be vetted through the HRPP *before* beginning. Further information concerning the HRPP Applicability Review Process is available from the Headquarters Marine Corps MF National Research Section: [https://ehqmc.usmc.mil/sites/family/mfq/national\\_research/default.aspx](https://ehqmc.usmc.mil/sites/family/mfq/national_research/default.aspx).

In addition, you should review the relevant DoD and Marine Corps guidance related to human subjects protection and survey administration, including

- Department of Defense Instruction 3216.02, *Protection of Human Subjects and Adherence to Ethical Standards in DoD-Supported Research*
- Marine Corps Order 3900.18, *Human Research Protection Program (HRPP)*
- Secretary of Navy Instruction 3900.39D, *Human Research Protection Program*
- Marine Corps Human Research Protection Program (HRPP) and Institutional Review Board website (<http://www.tecom.marines.mil/Resources/HRPP/>)
- Marine Corps Order 5300.18, *Marine Corps Survey Program*
- 45 Code of Federal Regulations 46.102.

With this guidance in mind, you can begin to think about the design of your own evaluation.

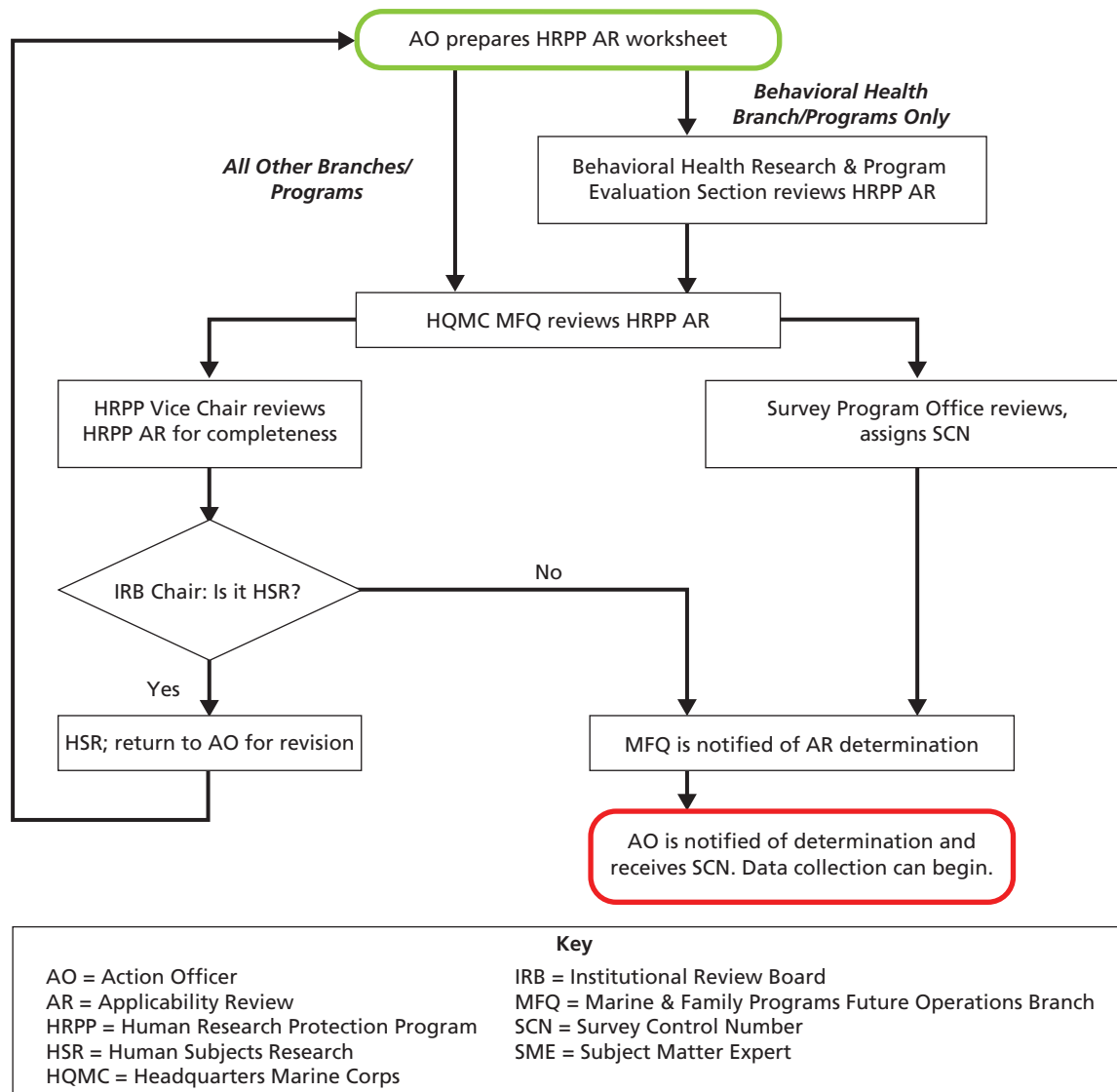
## Process Evaluation

Process evaluations generally aim to “[assess] what activities were implemented, the quality of the implementation, and the strengths and weaknesses of the implementation” (Chinman, Imm, and Wandersman, 2003). As shown in Figure 4.1, process evaluations correspond to the resources/inputs, activities, and output sections of the logic model and use measures of performance (MOPs) to assess program usage and implementation.

There are a number of factors that may be considered as part of a process evaluation:

- service utilization patterns, including participation, attendance, dosage (i.e., how much of a service was used), and/or attrition (Acosta et al., 2013; Centers for Disease Control and Prevention, 2011; Lalayants, 2012)

**Figure 4.2**  
**HRPP Applicability Review Process**



**NOTES:**

1. The AO (can be at the HQMC- or installation-level) submits appropriate HRPP AR worksheet to HQMC MFQ (National Research Section).
2. HQMC MFQ conducts a preliminary review of the HRPP AR worksheet and may involve SME input prior to submitting HRPP AR worksheet to the HRPP Vice Chair and any surveys to the Survey Program Office.
3. The HRPP Vice Chair and the Survey Program Office conduct concurrent reviews of the submitted HRPP AR worksheet.
  - a. HRPP Vice Chair makes preliminary determination regarding HSR and ensures all necessary documentation is present and forwards his or her determination to the IRB Chair.
  - b. Survey Program Office reviews survey, ensures scientific merit, and assigns SCN.
4. IRB Chair makes the final HSR determination based on submitted documentation from the HRPP Vice Chair
  - a. If determined to NOT be HSR, the original submitting AO is good to go but must have the SCN for any surveys prior to initiation.
  - b. If determined to be HSR, the original submitting AO will need to submit additional paperwork per policy.

- staff qualification and available resources (Centers for Disease Control and Prevention, 2011; W. K. Kellogg Foundation, 2004a)
- fidelity/adherence and quality of service implementation (Acosta et al., 2013; Chinman, Imm, and Wandersman, 2003)
- provider or client perceptions of the usefulness or relevance of the program (Acosta et al., 2013; Chinman, Imm, and Wandersman, 2003; Lalayants, 2012; McNeil, 2011)
- satisfaction with the program, including satisfaction of program staff and of individuals served by the program (Acosta et al., 2013; Chinman, Imm, and Wandersman, 2003; Lalayants, 2012).

This component of an evaluation helps you to understand how your PSC operated so as to meet goals and objectives. Well-designed process evaluations can produce critical findings on facilitators and barriers to implementation that offer important insights for future replication of a PSC.

Evaluating the implementation and usage of a PSC is important for two additional reasons. First, understanding how a program was implemented provides a context for the results of an outcome evaluation. If a program is not achieving the intended outcomes, it may be because it is not reaching the target population, it is not being implemented with fidelity, and/or program participants are dissatisfied with certain aspects of the program. Second, specific parameters for the implementation of PSCs in the MF portfolio have often been outlined in MCOs. For example, an MCO might specify that all active-duty Marines must receive a certain program or service, or describe eligibility criteria for a given PSC. An MCO might also provide guidance about how often a PSC's activities should take place. Therefore, a process evaluation will help you to determine whether your PSC is consistent with official guidance and whether there are areas for improvement.

### Identifying the Focus of the Evaluation

When designing a process evaluation, one key consideration is the specific aspects of implementation and usage that you are interested in assessing. Table 4.1 provides some examples of the types of questions that might guide your process evaluation to better understand the inputs, activities, and output components of your logic model. Note that this list is not intended to be exhaustive but provides a starting point for designing your process evaluation.

Using Worksheet 4.1, you can begin to identify the focus of your process evaluation. On the left, place a check mark next to the categories that you are interested in, and on the right, specify which aspects of each category you will measure. Examples based on our sample PSCs are presented following.

### Outcome Evaluation

Outcome evaluations focus on the short-term and intermediate outcomes expected to result from a PSC. Outcome evaluations “[attempt] to document whether or not the program caused an improvement among the participants on certain areas of interest . . . and by how much” (Chinman, Imm, and Wandersman, 2003). Rather than evaluating the implementation of a program, outcome evaluations focus on program results or outcomes (McNeil, 2011)—that is, whether changes that are expected as a result of the program actually occurred (Acosta et

**Table 4.1**  
**Questions That Guide Process Evaluation**

| Category                                                        | Sample Questions                                                                                                                                                                                                                                                                   |
|-----------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Staff qualifications and available resources (resources/inputs) | <ul style="list-style-type: none"> <li>• Did program staff have the appropriate credentials or certifications?</li> <li>• Were there adequate resources available for the program?</li> <li>• Does the PSC have information technology needed to support its functions?</li> </ul> |
| Fidelity of implementation (activities/outputs)                 | <ul style="list-style-type: none"> <li>• Was the program administered according to the curriculum?</li> <li>• Were ineligible participants appropriately screened out?</li> <li>• Was the activity offered the required number of times?</li> </ul>                                |
| Relevance or usefulness (activities/outputs)                    | <ul style="list-style-type: none"> <li>• Did the activities meet the interests and needs of participants?</li> </ul>                                                                                                                                                               |
| Usage (outputs)                                                 | <ul style="list-style-type: none"> <li>• How many people used the PSC?</li> <li>• Did the program reach its target audience?</li> <li>• Did the activities adequately meet demand?</li> <li>• How much of the service did people use?</li> </ul>                                   |
| Satisfaction (activities)                                       | <ul style="list-style-type: none"> <li>• How satisfied were participants?</li> </ul>                                                                                                                                                                                               |

al., 2013; Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury, 2012). Outcome evaluations aim to answer questions such as

- To what extent were the intended outcomes achieved?
- Did participants obtain the intended knowledge, skills, abilities, or behaviors?
- For whom were these outcomes achieved?
- During what time frame were outcomes met?
- Were outcomes sustained over the intermediate term?

These outcomes are often measured at the client level or program level (W. K. Kellogg Foundation, 2004a). For this toolkit, outcome evaluations correspond to the short and intermediate outcomes on your logic model.

There are several design options that can be used for an outcome evaluation. These options vary in their level of methodological rigor. In general, the more rigorous the design, the more confident you can be that any changes observed are the result of your PSC and not other factors (e.g., other PSCs, outside factors such as social support or the passage of time). We will review three outcome evaluation design options, as well as the benefits and challenges of each.

### Experimental Evaluations

Experimental evaluations, also known as **randomized controlled trials (RCTs)**, compare participants randomly assigned to participate in a PSC to individuals assigned to a control group.

The benefit of random assignment is that, generally speaking, it will produce two fairly equivalent groups (What Works Clearinghouse, 2014). This means that groups should be similar with respect to factors such as demographic characteristics (e.g., age, race, gender) or the characteristics that led them to be interested in the program (e.g., level of physical fitness, clinical characteristics, childcare or employment needs, relationship skills). Because the groups are equivalent, you can feel more confident that any outcomes observed among PSC participants and not among members of the control group are more likely due to participation in the PSC (Acosta et al., 2013). In fact, RCTs are considered the gold standard of research designs

**Worksheet 4.1****Identifying the Focus of Your Process Evaluation**

|                                                                                          |                                                                                                                                        |
|------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/> Staff qualifications and available resources (resources/inputs) | Specific aspects<br>(e.g., staff credentials/certifications, availability of resources)<br>• _____<br>• _____<br>• _____<br>• _____    |
| <input type="checkbox"/> Fidelity of implementation (activities/outputs)                 | Specific aspects<br>(e.g., adherence to curriculum, appropriate screening for eligibility)<br>• _____<br>• _____<br>• _____<br>• _____ |
| <input type="checkbox"/> Relevance or usefulness (activities/outputs)                    | Specific aspects<br>(e.g., activities matched participant needs and interests)<br>• _____<br>• _____<br>• _____<br>• _____             |
| <input type="checkbox"/> Usage (outputs)                                                 | Specific aspects<br>(e.g., attendance, demand met, target population served)<br>• _____<br>• _____<br>• _____<br>• _____               |
| <input type="checkbox"/> Satisfaction (outputs)                                          | Specific aspects<br>(e.g., among what participants, with which part of the PSC)<br>• _____<br>• _____<br>• _____<br>• _____            |
| <input type="checkbox"/> Other (outputs)                                                 | Specific aspects<br>• _____<br>• _____<br>• _____<br>• _____                                                                           |

because the results can determine the effectiveness of the program on the measured outcomes (National Center for Education Evaluation and Regional Assistance, 2003).

However, although they are the most methodologically rigorous option, RCTs can be challenging to implement in real-world settings. For example, you must have potential participants who are willing to possibly forgo the program or service because they have been assigned to the control group. There are ethical considerations related to this as well—what does it mean to withhold a PSC from individuals who may benefit? There are some alternate designs that help to address these ethical considerations, such as wait-list control groups, in which the PSC would be withheld from one group during the evaluation but members of that group would be offered the opportunity to participate once the evaluation was complete. However, for a wait-list control design, there still must be one group willing to wait to participate in your PSC activities. Finally, RCTs are more resource-intensive, as they require funding for recruitment of participants and assignment to groups. Costs can also be high because RCTs require enough resources to collect data from PSC participants *and* from the control group.

**Example Worksheet 4.1****Identifying the Focus of Your Process Evaluation**

|                                                                                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|-----------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input checked="" type="checkbox"/> Staff qualifications and available resources (resources/inputs) | Specific aspects <ul style="list-style-type: none"> <li>• <i>(L.I.N.K.S.)</i> How often was session content updated?</li> <li>• <i>(FMEAP)</i> Did all staff complete their required continuing education? How many continuing education programs did they attend?</li> <li>• <i>(Equipment Rental)</i> What equipment was available and in what quantities?</li> </ul>                                                                                       |
| <input checked="" type="checkbox"/> Fidelity of implementation (activities/outputs)                 | Specific aspects <ul style="list-style-type: none"> <li>• <i>(L.I.N.K.S.)</i> Did staff deliver L.I.N.K.S. workshops according to the set curriculum?</li> <li>• <i>(FMEAP)</i> Did career coaches appropriately assess for and identify the job-related goals of family members they served?</li> <li>• <i>(Equipment Rental)</i> Was the Equipment Rental open for all required hours?</li> </ul>                                                           |
| <input checked="" type="checkbox"/> Relevance or usefulness (activities/outputs)                    | Specific aspects <ul style="list-style-type: none"> <li>• <i>(L.I.N.K.S.)</i> Did the curriculum cover all of the topics that family members would need to navigate life in the Marine Corps?</li> <li>• <i>(FMEAP)</i> Did career training sessions cover common job search needs (e.g., resume writing, interviewing skills)?</li> <li>• <i>(Equipment Rental)</i> Did skill and safety lessons cover topics needed to safely operate equipment?</li> </ul> |
| <input checked="" type="checkbox"/> Usage (outputs)                                                 | Specific aspects <ul style="list-style-type: none"> <li>• <i>(L.I.N.K.S.)</i> Were all interested families able to enroll in a L.I.N.K.S. session?</li> <li>• <i>(FMEAP)</i> How many family members participated in career coaching?</li> <li>• <i>(Equipment Rental)</i> How many individuals and commands/units checked out equipment?</li> </ul>                                                                                                          |
| <input checked="" type="checkbox"/> Satisfaction (activities)                                       | Specific aspects <ul style="list-style-type: none"> <li>• <i>(L.I.N.K.S.)</i> What proportion of family members reported that they like the program?</li> <li>• <i>(FMEAP)</i> Did family members who attended sessions report that they were satisfied with the courses?</li> <li>• <i>(Equipment Rental)</i> Were Marines and dependents satisfied with the equipment they used?</li> </ul>                                                                 |

**Quasi-Experimental Evaluations**

Quasi-experimental evaluations compare PSC participants to individuals in a comparison group. Unlike experimental designs, however, the comparison group is not randomly assigned (Defense Centers of Excellence, 2012). Instead, a comparison group is selected to be as similar to the group participating in the PSC as possible. For example, if your PSC serves military families with children under the age of 10, you would identify a group of military families with children under 10 who did not participate in or use the PSC.

There are certain advantages to quasi-experimental evaluations. This type of evaluation is typically less logistically challenging to implement than an RCT because you do not have to worry about random assignment or withholding or postponing services from one group. The latter also means the ethical dilemma posed by RCTs is less of a concern. In addition, because there is a comparison group, this design can still help evaluators understand differences between individuals who participate in an activity and those who do not (Defense Centers of Excellence, 2012).

There are disadvantages to quasi-experimental designs, though. Because participants have not been randomly assigned, the comparison group in a quasi-experimental evaluation cannot

be assumed to be equivalent (Martin et al., 2014). The comparison group may be different from PSC participants in some key ways—and these differences may have contributed to the decision not to participate in the PSC. In the example of families with young children, it may be that their children are older on average, or that they live farther from the military installation, which affected their decision to participate. This is important, because these same factors may also contribute to the outcomes that are being measured. For example, if you measure quality of Marine Corps life as an outcome, families that participate in the program may rate their quality of life higher. But is this because they were a part of the program, or because they live closer to base and therefore were able to take advantage of on-base programs?

Another disadvantage is that it can be difficult to identify a matched group for the purpose of comparison. Generally, when you conduct a quasi-experimental evaluation, you would identify specific characteristics that you want to be similar between the group using your PSC and comparison group. However, there are always trade-offs across characteristics—for example, it may be easy to find a group that matches on gender and race/ethnicity, but not on age. Finally, similar to RCTs, costs can be high because data are collected from PSC participants and from members of the comparison group (Acosta et al., 2013).

### **Observational Evaluations**

Observational evaluations focus on participants of a PSC alone without a control or comparison group. These types of evaluations can take a couple of different formats (Chinman, Imm, and Wandersman, 2003). In some observational evaluations, participants are administered surveys before and after participation in a PSC to examine changes in knowledge, attitudes, skills, or behaviors associated with participation. Other observational evaluations may only administer post-participation measures, but ask participants to reflect on changes that occurred while they were participating in the program or using the services (Acosta et al., 2013).

A key benefit of observational evaluations is that they are less resource-intensive (Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury, 2012). Data are collected only from individuals who participated in the PSC, and this can be a simple way to measure change that happened over the course of participation.

However, there are disadvantages as well. First, because there is no control or comparison group, it is not possible to rule out other explanations for the outcomes observed (Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury, 2012). This challenge especially applies to observational studies that only administer post-participation measures and ask participants to reflect on changes (Acosta et al., 2013). Participants may have difficulty recalling what took place before they participated in a PSC, or their memory may have been affected by the passage of time, or by their participation.

The advantages and disadvantages of the outcome evaluation methods just discussed are summarized in Table 4.2.

### **Types of Evaluation Data**

Another important aspect of designing an evaluation is deciding what type of data you will use. In the next section, we describe two considerations related to data sources: (1) the distinction between qualitative and quantitative methods, and (2) the difference between primary and secondary data.

**Table 4.2**  
**Advantages and Disadvantages of Outcome Evaluation Design Options**

| Evaluation Design                                                                                                     | Advantages                                                                                                                                                                                                                                                         | Disadvantages                                                                                                                                                                                                                                                                                                                                        |
|-----------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Experimental<br>Compares program participants to randomly assigned control group                                      | <ul style="list-style-type: none"> <li>• “Gold standard” evaluation design for program effectiveness evaluation</li> <li>• Strong control groups that are equivalent to treatment groups making results clearer</li> <li>• Builds PSC research capacity</li> </ul> | <ul style="list-style-type: none"> <li>• Difficult to implement in real-world settings</li> <li>• High level of resources and expertise needed for recruitment, retention, and data collection</li> <li>• Reluctance of some individuals to participate in RCTs</li> <li>• Reduces the number of individuals served if recruitment is low</li> </ul> |
| Quasi-experimental<br>Compares program participants to comparison group                                               | <ul style="list-style-type: none"> <li>• Comparison group drawn from the same or similar target population, so similar to treatment group</li> <li>• Able to serve more individuals, since not losing half to randomization</li> </ul>                             | <ul style="list-style-type: none"> <li>• Difficulty recruiting and retaining individuals in a comparison group</li> <li>• Need for sophisticated statistical adjustments that may or may not adequately control for bias because of differences between groups</li> <li>• Funding and staffing issues of concern for feasibility</li> </ul>          |
| Observational—pre-post design<br>Examines program participants alone, comparing status before and after participation | <ul style="list-style-type: none"> <li>• Less resource-intensive, as no data collection from control or comparison group required</li> <li>• Straightforward way to measure change</li> </ul>                                                                      | <ul style="list-style-type: none"> <li>• More challenging to rule out alternative explanations for effects that are observed</li> </ul>                                                                                                                                                                                                              |
| Observational—posttest only<br>Examines program participants alone, using post-participation assessment only          | <ul style="list-style-type: none"> <li>• Less resource-intensive</li> <li>• Only one assessment point needed</li> </ul>                                                                                                                                            | <ul style="list-style-type: none"> <li>• More challenging to rule out alternative explanations for effects that are observed</li> <li>• Recall bias or inaccuracies when asked to report on preactivity status</li> </ul>                                                                                                                            |

SOURCE: Acosta et al., 2013; RAND NDRI review of relevant literature.

### Objective and Subjective Data

Both objective and subjective data may be collected as part of a program evaluation (Hourcade, Parette, and Anderson, 2003). *Objective data* refers to information that is observable and not subject to influence by an individual’s perceptions or beliefs. By contrast, *subjective data* capture personal opinions, beliefs, and experiences. Both types of data can be used for process and outcome evaluation. For process evaluation, objective data might include the number of people participating in a workshop, and subjective data could include ratings of satisfaction. For outcome evaluations, objective data could include scores on a knowledge test, while subjective data could be perceptions of social support or community connectedness. The selection of the most appropriate type of data will depend, in part, on the measures of performance and effectiveness identified for the evaluation.

### Qualitative and Quantitative Data

Another consideration is whether your evaluation questions are best answered with qualitative data, quantitative data, or some combination.

*Qualitative data* include descriptive information about the PSC and the implementation process and are typically associated with the process evaluation component of the overall evaluation. To collect qualitative data, evaluators may conduct individual and group interviews

with key informants (e.g., program staff and administrators, program participants), conduct focus groups, engage in observation of an activity, or ask open-ended questions on surveys (Chinman, Imm, and Wandersman, 2003; LeBlanc and Lacey, 2002; W. K. Kellogg Foundation, 2004a; World Health Organization, 2013). These kinds of qualitative data are often reviewed and analyzed to identify major themes that arise across interviews and informants. Qualitative methods can provide rich information or context. For example, focus groups with program participants may be used to understand the aspects of a PSC that were most relevant or useful, and whether there were certain aspects of the method of delivery that made it especially useful. Observations can be used to determine whether a PSC with a specific curriculum has been implemented with fidelity to the curriculum. However, these methods can also be time-consuming and, therefore, resource-intensive.

*Quantitative data* are in a numerical format and can contribute to both the process and outcome evaluation components of the evaluation (World Health Organization, 2013). For example, this could include quantities—such as numbers of people served by a PSC or responses on structured knowledge questions about a topic—that produce a score on a numerical scale. These data are often collected through surveys or questionnaires or may be obtained from existing clinical or administrative data (Chinman, Imm, and Wandersman, 2003; Lalayants, 2012; World Health Organization, 2013). Quantitative data can be analyzed using traditional statistical methods that allow for comparisons between groups or assess change over time. Quantitative data can be used as part of a process evaluation (e.g., satisfaction ratings made on a numerical scale) and are generally used to measure outcomes.

Other sources of qualitative and quantitative data may include reviews of existing reports or documents, existing monitoring or evaluation systems, and expert panels (W. K. Kellogg Foundation, 2004a). Data can be collected across multiple levels, including the individual, group, and program level (LeBlanc and Lacey, 2002; Unger et al., 2000).

For this toolkit, we focus on the use of quantitative data. In part, this is because quantitative data are often more amenable to being aggregated across large groups of people and analyzed, especially when limited resources are available. However, it is useful to keep in mind that qualitative data may also be used to answer your evaluation questions, particularly questions related to your process evaluation.

Both qualitative and quantitative data can be either objective or subjective. Table 4.3 provides examples of qualitative, quantitative, objective, and subjective data that may be relevant to your evaluation.

### **Primary and Secondary Data**

For quantitative methods, another consideration is the use of primary data, secondary data, or some combination (World Bank, 2009).

*Primary data* refers to information being collected specifically for the evaluation. Primary data sources relevant to a process evaluation might include attendance lists or satisfaction surveys administered to PSC participants, and primary data sources relevant to outcome evaluation might include behavioral health measures or knowledge tests. Benefits of primary data are that the PSC and/or evaluator determines what is collected and when and is aware of the limitations of the data. However, the burden of data collection falls on the PSC, which can present a challenge if you have few resources (including staff and funding) or serve a large volume of participants.

**Table 4.3**  
**Examples of Qualitative, Quantitative, Objective, and Subjective Data**

|                   | Objective Data                                                                                                                  | Subjective Data                                                                                                                                      |
|-------------------|---------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|
| Qualitative data  | <ul style="list-style-type: none"> <li>• Observer ratings of workshop leaders to rate fidelity to program curriculum</li> </ul> | <ul style="list-style-type: none"> <li>• Focus group perceptions of relevance of program curriculum</li> </ul>                                       |
| Quantitative data | <ul style="list-style-type: none"> <li>• Number of workshop participants</li> <li>• Number of positive drug tests</li> </ul>    | <ul style="list-style-type: none"> <li>• Participant ratings of satisfaction with a PSC</li> <li>• Participant ratings of quality of life</li> </ul> |

*Secondary data* refers to data compiled for purposes outside the administration of the PSC, but which are relevant to the outcomes of the PSC. For example, the Marine Corps Quality of Life Survey is administered periodically to measure perceptions of and satisfaction with aspects of the USMC that contribute to quality of life. Some of these questions may be relevant to your process evaluation (e.g., items that assess satisfaction with specific USMC services) or outcome evaluation (e.g., ratings of health, quality of life). Another source of secondary data might be data collected as part of the evaluation of another PSC that reaches a similar audience, has similar goals, or has similar short-term and intermediate outcomes.

One important benefit of secondary data is that the burden for data collection falls outside the PSC, which can be helpful if a PSC serves many people or has few resources. In some cases, secondary data do not involve information collection from individual program participants: missed duty days, unit readiness indicators, and disciplinary infractions are examples of secondary data that may be useful for certain PSCs. However, in most cases, you do not have control over the content of the data (e.g., specific questions asked or numbers tracked). It can also be challenging to fully know the limitations to the data (e.g., mistakes or errors in the data) (Hox and Boeijs, 2005; World Bank, 2009). Additional considerations related to primary versus secondary data are discussed in the next chapter.

## Evaluation Considerations

Several factors influence evaluation design. In this section, we introduce a number of considerations that may guide your decisions as you design your process and/or outcome evaluation. For outcome evaluations, there are considerations related to the number of participants and the availability of comparison groups. Several other considerations pertain to both process *and* outcome evaluations, including the timing of evaluation and available evaluation resources, such as the infrastructure for data collection, human capital needs, and financial demands. Other considerations for both process and outcome evaluations include data security and the time frame of the evaluation.

Evaluations often take place in situations in which there are certain constraints on resources, expertise, participants, or infrastructure. These are important to take into consideration when selecting your evaluation design. Contextual considerations are important as well. For example, if your PSC serves Marines preparing to deploy, you may need to consider the way this affects your ability to collect follow-up data. This section will help you to review

your options and ultimately select an evaluation design that is as rigorous as possible given the capacity of your PSC.

After reviewing this section, use Worksheet 4.2 to take notes about your activity that may be relevant to each of these factors.

### **Availability of Control or Comparison Groups**

The availability of a control or comparison group is an issue that largely pertains to outcome evaluation. As described, the most rigorous outcome evaluation designs involve the use of a control or comparison group. However, it may be difficult to identify an appropriate group. If you are implementing an RCT, the recruitment of a control group can be challenging because you need staffers who can identify and recruit potential study participants and be able to explain the benefits of participating in the study. You also need to find a group of individuals who are interested in the activity but understand that by participating in the study they will not be able to participate in the activities or receive the program or service.

As noted earlier, quasi-experimental designs circumvent this challenge through the identification of a comparison group that is not receiving services but is as similar to the program participants as possible. In some cases, it may not be too difficult to identify and reach a comparison group. For example, maybe you have people who are interested in your PSC, but not enough slots for them to participate at this time. Or maybe you have enough capacity to serve these individuals, but your program or service does not work with their schedules, so they are unable to participate at this time. This might be a natural comparison group.

If there is not a readily available comparison group, though, you will have to determine whether an appropriate comparison group exists and how difficult it might be to reach those individuals. Keep in mind that, ideally, members of the comparison group are as similar to program participants as possible. What are the key characteristics that you want to match the comparison group on? Is there a way to systematically reach these individuals, or would recruitment be a time- or resource-intensive process? These factors may influence your evaluation design.

### **Number of Participants**

Another important consideration is the number of participants that you will be able to include in your evaluation. This issue is relevant to both process and outcome evaluations. An initial consideration is whether you will include the entire population of individuals who use your PSC, or a sample or subset of the population. In an ideal world, you may want to survey the entire population of individuals who use your PSC. There are many benefits to surveying the population (Daugherty et al., 2016):

- **Full representation:** By including the entire population, you can be sure that you have not excluded unique cases or subsets of individuals who use the program.
- **Statistical precision:** Larger numbers of observations and greater coverage of the population will ensure that you are measuring characteristics of the population and changes that occur more precisely.
- **Fostering buy-in:** Including all participants of your PSC in an evaluation may help them to feel better integrated into your PSC, more invested in the evaluation process, and that their input and voices matter.

However, including the population of participants in your evaluation can also be time- and resource-intensive. For example, if your PSC serves a large number of participants, it may be too burdensome to collect evaluation data from all individuals using the activity. For this reason, you may choose to select a subset of participants for your evaluation. Some reasons to select a sample (Daugherty et al., 2016) are:

- **Limited resources:** When you serve a large population, it can be resource-intensive to survey everyone, including costs of survey materials, time to collect and analyze data, and time needed for recruitment and outreach efforts. Using a sample helps to mitigate these costs.
- **Survey fatigue:** Surveys are a common part of military life, and Marines (and potentially family members) may receive surveys from a variety of sources. Overburdening individuals with surveys can lead to a lower response rate across all surveys. Therefore, you may consider using a sample to limit survey fatigue.
- **Opportunity costs:** When individuals complete surveys, they may be contributing less to other duties, including work responsibilities. This may be a particularly important consideration if your PSC serves Marines, not just family members.

These considerations may help you to determine whether your evaluation will focus on the entire population using your PSC, or if you will identify a sample.

There are a few additional considerations related to number of participants. If you decide to select a sample to participate in your evaluation, one overarching consideration is that, in general, a larger sample will allow you to have more confidence that your sample is representative of the population using your PSC.<sup>3</sup> Larger samples also allow you to detect smaller effects between treatment groups; that is, between those individuals who do and do not participate in the program. This is true for both process and outcome evaluations. However, the number of participants in your PSC or the flow of participants through your PSC may constrain the number of people who can be included in the evaluation, even if you collect data from all participants. On a related note, if you choose to use qualitative methods, such as interviews or focus groups, it can be time-consuming to have a large sample: even if you limit interviews or focus groups to 30 to 60 minutes, the time needed adds up quickly with the number of participants who are included.

There are also particular considerations related to outcome evaluations. For example, the number of participants that you need to include will depend on factors such as the design of your outcome evaluation and the degree of change expected on the outcome of interest. All other factors held constant, RCTs require smaller sample sizes than quasi-experimental designs because the two groups are more equivalent and, thus, more comparable. Therefore, a smaller sample is needed to detect the true effect of the PSC. Similarly, if you expect a large change on the outcomes measured, a smaller sample size is needed for the analysis.

Because there are a number of factors that can influence your ability to detect a significant effect (e.g., evaluation design, strength of effect, sample size), a statistician can help you to

---

<sup>3</sup> A large sample will not always result in an unbiased sample. For example, a sample of 1,000 spouses of junior enlisted Marines would provide very precise estimates on any given outcome measure. Those estimates, however, would tell you something only about that particular group of spouses, and the results may not hold for spouses of senior enlisted personnel or junior officers.

determine an adequate number of participants for your evaluation, if you have access to one. If not, then it is generally good practice to include as many participants as possible in your evaluation, within the constraints of your PSC and available resources. If you are unable to obtain a large sample for your evaluation, you should keep in mind that your sample may differ from the larger population that uses the PSC in some way (e.g., demographic factors, frequency with which they use the program, or level of enthusiasm for the services). Therefore, if you are unable to obtain a large sample for your evaluation, this limitation should be kept in mind when you interpret the results.

There are ways that you can increase the representativeness of your sample if you are unable to collect data from all participants in your PSC (Teddlie and Yu, 2007). One option is to select a *random sample*. Random sampling means that each person who uses your PSC during the evaluation period has an equal chance to be enrolled in the evaluation. For example, maybe you decide that you are going to invite one-third of your PSC participants to take part in the evaluation. You could use a random number generator (e.g., random.org) to create a randomly ordered list of the digits 1, 2, and 3. As each person is referred to the program, you would move to the next number on the list, and you could decide that everyone matched with number 1 is going to be invited to be part of the evaluation.

Another option you might consider is *stratified sampling* (Teddlie and Yu, 2007). Stratified sampling helps you to select a sample that is similar to your overall population with respect to certain characteristics. To use this method, you would first identify characteristics of the population especially important to your program—for example, maybe you want to be sure that you collect data from spouses of both enlisted service members and officers, or you want to collect data from both male and female Marines. Then, you would select participants from each of those groups to ensure that they were represented in your sample. For example, if you want to select spouses of both officers and enlisted service members, you could randomly select participants from all spouses of officers enrolled in the PSC *and* randomly select participants from all spouses of enlisted service members enrolled. This will ensure that you have people from both groups participating in the evaluation. You can also take stratified sampling a step further to make sure that your sample is similar to the overall population of PSC participants. To do this, you would determine the proportion of individuals in each group that your PSC serves. For example, maybe 80 percent of the participants your PSC serves are spouses of enlisted service members, and 20 percent are spouses of officers. Then, you would aim to match these proportions with your sample. So, if you had only the resources to collect data from 100 participants, you would aim for 80 of those to be spouses of enlisted service members (80 percent), and 20 to be spouses of officers (20 percent).

In Chapter Six, we revisit this topic and provide tips for increasing the response rate for your evaluation.

### Timing of Evaluation

Timing is a factor relevant to both process and outcome evaluation. One consideration is when to collect data as part of your evaluation. Ideally, an evaluation is planned and implemented prior to the start of the PSC. This has several advantages for process and outcome evaluations. Regarding your process evaluation, this gives you the opportunity to collect usage data in real time, rather than having to find ways to measure these elements after the fact. Many aspects of a process evaluation are well suited to being collected continuously over the course of an evaluation, such as determining whether staff members have appropriate qualifications or whether a

workshop is implemented with fidelity. That said, other components of your process evaluation might be implemented after participants have participated in your PSC, such as satisfaction surveys.

For your outcome evaluation, implementing an evaluation prior to the start of the PSC gives you the opportunity to collect pretest data from your participants. For example, with the L.I.N.K.S. program, you would administer your evaluation measures before the workshop begins and then again after it has ended. Another consideration related to outcome evaluation is the timing of outcomes. Some effects are observed in the short term (e.g., immediately after the program ends), whereas others will not be observed until sometime after the program has ended (e.g., six months later).

Timing of the evaluation can become more complex when you have a service or capability, rather than a more formal program. For example, one individual could use Outdoor Recreation Equipment Checkout and Equipment Rental multiple times sporadically over the course of several months. What is the best time to survey this person—after each equipment rental? At set time intervals, such as every month? These decisions are made more challenging by the fact that you cannot predict when a given person will decide to use the capability. In this case, you might decide that it makes the most sense to survey a person every time they rent equipment, or that it is simpler to email all individuals who rented equipment each month a link to a survey at the end of the month. Timing of equipment checkout or rental may also vary across seasons because many recreational activities are weather dependent.

Finally, a consideration that may impact the timing of your evaluation is whether there are specified reporting standards for your activity. For example, the MCOs for some PSCs indicate that a certain number of activities must take place in a six-month or one-year period. In this case, data would have to be collected often enough to evaluate the PSC on that schedule. This issue generally pertains to process evaluation.

On a related point, it can be useful to consider whether the evaluation is being implemented to determine whether a PSC has achieved a specific target, metric, or benchmark or is intended to be part of a larger cycle of continuous improvement. These considerations may also inform the timing of the evaluation (for more detail about continuous quality improvement and how this may apply to your PSC, see Chapter Seven).

### **Evaluation Expertise**

As you plan for your process and/or outcome evaluation, consider who will actually conduct the evaluation and whether they have any knowledge or experience with program evaluation. Evaluation expertise is helpful if you need to design data tracking forms or develop measures for your evaluation (as discussed in more detail in Chapter Five). Also, if you will be using existing staff within your PSC who are less familiar with program evaluation, it may be more challenging to implement the more technical evaluation designs, such as RCTs. Another option to consider is whether your PSC has established relationships with any academic institutions, community-based organizations, or contractors with evaluation expertise, or whether there are staff in your branch's headquarters office who could assist in the design and execution of the evaluation. If so, a more rigorous evaluation design or more involved data collection efforts could be feasible.

In general, when thinking about the personnel needed for an evaluation, you will likely need people with expertise in the following domains:<sup>4</sup>

- **PSC content:** This person is familiar with the PSC, its objectives and activities, and the day-to-day operation. This person will be an asset when developing the logic model, determining what type of evaluation plan will be feasible to implement, and in ultimately interpreting results of the evaluation.
- **Evaluation design and methods:** This person will design the evaluation plan and oversee the implementation of the evaluation. As mentioned, this person may be internal or external to the PSC being evaluated, depending on the expertise of PSC staff. Sometimes, it can be helpful to have an external evaluator to provide a more objective perspective, though internal evaluators may be more familiar with the inner workings, culture, and staff of the PSC and broader military installation.
- **Data analysis:** Sometimes, evaluation personnel responsible for designing the evaluation plan will also have the expertise necessary to analyze and interpret the data yielded by the evaluation. However, it can also be useful to have team members whose primary role is to plan and conduct data analysis. Multiple data analysis experts may be needed, depending on the type of data collected (e.g., qualitative versus quantitative data).

In addition, it can be helpful to have someone in a position of authority who can serve in a decisionmaking capacity. This is a person who can review the results of an evaluation and has the authority needed to make any changes recommended.

### Available Evaluation Resources

Program evaluations require time, infrastructure, and other resources, and certain types of outcome evaluations (e.g., RCTs) can be more resource-intensive than others. As you design your evaluation, it will be important to consider the resources available. Key evaluation resources include

- **Infrastructure:** Is there an existing infrastructure to collect data? This may depend on the ways that data are collected (e.g., interviews, surveys, attendance logs), the way that data are stored (e.g., paper files, computerized records), and the resources needed to analyze data (e.g., availability of spreadsheet or analytic software, capacity to review interview transcripts).
- **Human capital needs:** Evaluation expertise is one important factor, as described above, but so is manpower. Does your PSC have enough staff to carry out the evaluation, or do your staff members have time in their schedules to implement the evaluation plan? Consider all stages of the evaluation, from design to data collection to analysis. Keep in mind that data collection can be time-consuming, particularly if you need staff to track attendance, administer satisfaction surveys, and/or administer outcome measures.
- **Financial demands:** In part, financial considerations are related to human capital needs—do you have the staff that you need, and can they spare the time necessary for the evaluation? Or do you need to bring additional staff on board? Data collection may also

---

<sup>4</sup> For more details about identifying an evaluation team and other relevant stakeholders, see W. K. Kellogg Foundation (2004a).

have costs involved. One thing to keep in mind is that keeping the costs of data collection low may increase the human capital costs and vice versa. For example, you may be able to keep data collection costs low by administering pencil-and-paper surveys but then need staff members to enter this information into a computerized format for analysis. By contrast, you may be able to automate the data collection process using online surveys or computerized registration pages, but these systems may require more of an up-front cost.

There are ways to balance these considerations when selecting an evaluation design. Regarding the design of your process evaluation, if you have limited resources, you may consider using as many existing sources of data as possible, such as online registration data. In terms of outcome evaluation, if you have limited resources, you may consider one of the less resource-intensive outcome evaluation methods (e.g., observational) or explore the use of secondary data.

### **Data Security**

Before undertaking an evaluation of your PSC, it is important to establish a data security plan. This is essential for both process and outcome evaluation. Some of the information that you are collecting about participants could be sensitive (e.g., mental health symptoms, financial standing, marital satisfaction). Therefore, it is critical to have a plan to safeguard the identity of your participants and any information collected about them. If you are administering measures only once (e.g., as part of an observational study), it may be possible to collect your data without any identifying information, such as names or dates. However, if you are planning to collect data at multiple time points and be able to link it to a specific individual, you will have to collect some identifying data.

An additional challenge is introduced if your program does not typically collect identifying information from participants. You will have to balance considerations related to the importance of certain pieces of data with the need to request identifying information from your participants. Some individuals may be reluctant to provide their names or contact information if it would not otherwise be required. For this reason, it is preferable that participation in the evaluation be voluntary. This will help individuals interested in your PSC feel comfortable accessing your services or activities while not participating in the evaluation.

There may be data security concerns related to your method of data collection. For example, if you are sending a link to a survey via email, you will need the participants' email addresses and will also have to consider whether the survey software is collecting personal data, like IP addresses. If data are stored electronically on a computer linked to a USMC or other server, you will also have to consider who has access to the data and whether it is secure from outside users.

No matter how you collect data from participants, you must also address informed consent. Informed consent is the process by which participants understand the risks and benefits associated with providing information used for program evaluation. Typically, risks are low. However, there may be instances (e.g., substance use treatment) where inadvertent disclosure could cause emotional distress to the service member or his or her family. Informed consent is designed to make sure that participants fully understand the risks associated with their participation in a data collection effort. If participants do not believe that the benefits outweigh the risks, they may decide not to participate.

### Evaluating Contracted Services

Certain complications may arise if your PSC uses a contractor to provide any of its activities. If this is the case, you will first have to decide whether your evaluation will focus on just those activities carried out by the USMC, or whether you also want to evaluate the services provided by the contractor. If you want to evaluate the services provided by the contractor, keep these questions in mind:

- What level of oversight does the USMC have over the contractor? Does it carry out most activities for your PSC, or just a small aspect?
- Is there a liaison to the contractor, or how frequently does your PSC communicate with the contractor?
- Does it introduce additional data security concerns to request information from the contractor (e.g., asking it to provide identifiable information that is otherwise hosted on a secure external server)?
- Does the contractor have staff available to assist with the evaluation efforts, or will the tasks and/or costs be the responsibility of the USMC?

Above all, in this situation, it will be important to include the contracting organization in discussions about evaluation design before the evaluation begins. This will ensure that your evaluation design is feasible for your PSC and the contractor, and that it will provide you with meaningful process and outcome information.

### Selecting an Evaluation Design

Now, based on your answers to the questions in Worksheet 4.2, begin to think about your evaluation design.

Use Worksheet 4.3 to help plan the method that you will use to implement your evaluation. In **Section 1**, indicate the anticipated start and end dates of your evaluation. The start date should correspond to the date that you will begin to implement evaluation-related activities (e.g., when data collection begins). Start at the very beginning, with the creation of data collection systems and other preliminary activities needed before you can start collecting data. Then consider how long you want to collect data for both the process and outcome evaluation. After data collection is complete, build in time for analysis and reporting of results. Select an end date that corresponds to the completion of all evaluation activities, not just data collection (e.g., writing a final report). This will ensure that you have sufficient time for your evaluation. Be realistic in your estimates of timing. In fact, it can be helpful to build in a cushion of time, because some evaluation activities may take longer than expected.

In **Section 2**, you will plan for your methods and data sources. The first section focuses on the process component of the evaluation. First, you will indicate each data source/method that you will use, such as attendance records, focus groups, or satisfaction surveys. To some extent, the data sources that you collect will be dictated by the measures of performance and effectiveness that you have identified for your program and by how you choose to define those measures. The selection and definition of measures is covered in Chapter Five, so you may want to return to this template after you have read Chapter Five and completed the worksheets in that chapter as well. Next, indicate whether those data sources are already available to you,

**Worksheet 4.2****Evaluation Design Issues**

Now that you are familiar with some of the issues that may shape your evaluation design, complete the worksheet below. Consider the questions in each section and how they apply to the activity or set of activities that you would like to evaluate for your PSC. Your responses will be useful as you design your evaluation.

**Availability of Comparison Group**

- Would it be possible to implement a control group that does not participate in this activity?
- If you identified an existing comparison group, what characteristics would you want the group to have (i.e., how should it be similar to the population using your activity)?

Your notes:

---

---

---

---

---

---

**Number of Participants**

- Will you be using qualitative methods for your process evaluation?
- What is the outcome evaluation design for this activity?
- What degree of change in the outcomes is expected?
- Will you be able to include the population of participants, or will you select a sample (i.e., subset) from this population?
- How many individuals make use of this activity? Over what time frame?
- Do you have access to someone with the statistical skills to help you select the ideal sample size?

Your notes:

---

---

---

---

---

---

**Timing of Evaluation**

- How long is this activity? If not a set length, how long are participants typically engaged?
- Over what time frame do you expect to observe the short-term effects of this activity?
- Over what time frame do you expect to observe the intermediate effects of this activity?
- On what type of schedule is it possible to collect data? For example, can you collect data (a) before and after a person has participated, (b) only after a person has participated, or (c) on a set schedule (e.g., every 3 months)? Remember to consider the measurement of both process and outcome measures.
- Are there specified reporting standards for this activity (e.g., as outlined in an MCO)?

Your notes:

---

---

---

---

---

---

**Evaluation Expertise**

- Do PSC staff have evaluation expertise?
- Are there headquarters staff with relevant evaluation expertise?
- Are there other organizations (e.g., academic institutions, community-based organizations) that could assist with an evaluation effort?

Your notes:

---

---

---

---

---

---

|                                                                                                                                                                                                                                                                                                                                                                                |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Available Evaluation Resources</b> <ul style="list-style-type: none"> <li>• What staffing resources does your PSC have available to support the evaluation?</li> <li>• What financial resources are available?</li> <li>• What infrastructure is available (e.g., for data collection)?</li> </ul>                                                                          |
| Your notes:<br><hr/> <hr/> <hr/> <hr/> <hr/>                                                                                                                                                                                                                                                                                                                                   |
| <b>Data Security</b> <ul style="list-style-type: none"> <li>• What type of data would you like to collect? Is any of the information particularly sensitive?</li> <li>• Would you need to collect data over multiple time points, or be able to link responses to a specific person for any other reason?</li> </ul>                                                           |
| Your notes:<br><hr/> <hr/> <hr/> <hr/> <hr/>                                                                                                                                                                                                                                                                                                                                   |
| <b>Evaluating Contracted Services</b> <ul style="list-style-type: none"> <li>• What oversight do you have over the contractor, and what activities is it responsible for?</li> <li>• What data security concerns arise?</li> <li>• Who will be responsible for evaluation activities, and how will information be communicated between your PSC and the contractor?</li> </ul> |
| Your notes:<br><hr/> <hr/> <hr/> <hr/> <hr/>                                                                                                                                                                                                                                                                                                                                   |

or whether the collection method will have to be developed and implemented. After you have identified these data sources/methods, the next step is to indicate the timing of data collection for each data source/method. Note that for your process evaluation, many measures may be continuously monitored over the course of the evaluation.

The second portion of Section 2 focuses on your outcome evaluation. You will again identify your data sources/methods, availability, and schedule for data collection. Your outcome evaluation measures are more likely to be anchored to distinct phases of participation in the PSC, which is reflected in the worksheet.

In **Section 3**, you will provide information about the sample that corresponds to each of your sources, including a description of the sample and the expected sample size. Your sample may vary based on the specific data source or measure, and this worksheet allows you to plan for these variations. In turn, this information will be useful in completing **Section 4**, which allows you to plan for the resources needed for the evaluation, including a staffing plan, a consideration of the infrastructure needs you might have, and any additional costs associated with your evaluation.

Include as much detail as possible, as this will make the process of implementing your evaluation plan easier. Sample evaluation design plans for each of the three example PSCs are provided below. Note that these examples are fictional (e.g., number of participants, staff names) though based on the actual services and activities offered by each PSC. Although they

**Worksheet 4.3**  
**Designing Your Evaluation**

| Section 1: Time Frame                                                                                             |                          |                             |                           |                            |
|-------------------------------------------------------------------------------------------------------------------|--------------------------|-----------------------------|---------------------------|----------------------------|
| Expected start date: _____                                                                                        |                          |                             |                           |                            |
| Expected end date: _____                                                                                          |                          |                             |                           |                            |
| Section 2: Methods/Data Sources                                                                                   |                          |                             |                           |                            |
| Process Evaluation                                                                                                |                          |                             |                           |                            |
| <b>Data Source/Method</b><br>(e.g., attendance records, focus groups,<br>personnel records, satisfaction surveys) | Available now            | To be implemented           |                           |                            |
| _____                                                                                                             | <input type="checkbox"/> | <input type="checkbox"/>    |                           |                            |
| _____                                                                                                             | <input type="checkbox"/> | <input type="checkbox"/>    |                           |                            |
| _____                                                                                                             | <input type="checkbox"/> | <input type="checkbox"/>    |                           |                            |
| _____                                                                                                             | <input type="checkbox"/> | <input type="checkbox"/>    |                           |                            |
| _____                                                                                                             | <input type="checkbox"/> | <input type="checkbox"/>    |                           |                            |
| _____                                                                                                             | <input type="checkbox"/> | <input type="checkbox"/>    |                           |                            |
| Data collection schedule:                                                                                         |                          |                             |                           |                            |
| <b>Data Source/Method</b>                                                                                         | Timing                   |                             |                           |                            |
| _____                                                                                                             | _____                    |                             |                           |                            |
| _____                                                                                                             | _____                    |                             |                           |                            |
| _____                                                                                                             | _____                    |                             |                           |                            |
| _____                                                                                                             | _____                    |                             |                           |                            |
| _____                                                                                                             | _____                    |                             |                           |                            |
| Outcome Evaluation                                                                                                |                          |                             |                           |                            |
| <b>Data Source/Method</b><br>(e.g., knowledge tests, surveys of attitudes or<br>skills, behavioral measures)      | Available now            | To be implemented           |                           |                            |
| _____                                                                                                             | <input type="checkbox"/> | <input type="checkbox"/>    |                           |                            |
| _____                                                                                                             | <input type="checkbox"/> | <input type="checkbox"/>    |                           |                            |
| _____                                                                                                             | <input type="checkbox"/> | <input type="checkbox"/>    |                           |                            |
| _____                                                                                                             | <input type="checkbox"/> | <input type="checkbox"/>    |                           |                            |
| _____                                                                                                             | <input type="checkbox"/> | <input type="checkbox"/>    |                           |                            |
| _____                                                                                                             | <input type="checkbox"/> | <input type="checkbox"/>    |                           |                            |
| Data collection schedule                                                                                          |                          |                             |                           |                            |
| <b>Data Source/Method</b>                                                                                         | <b>Pre-Participation</b> | <b>During Participation</b> | <b>Post-Participation</b> | <b>Follow-Up (specify)</b> |
| <i>Example:<br/>Knowledge test</i>                                                                                | X                        |                             | X                         | X (6 months later)         |
| _____                                                                                                             |                          |                             |                           |                            |
| _____                                                                                                             |                          |                             |                           |                            |
| _____                                                                                                             |                          |                             |                           |                            |
| _____                                                                                                             |                          |                             |                           |                            |

| Section 3: Selecting/Describing Sample                                                                                                                    |                    |                                 |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------|---------------------------------|
| Process Evaluation                                                                                                                                        |                    |                                 |
| Data Source/Method                                                                                                                                        | Sample Description | Expected Sample Size            |
|                                                                                                                                                           |                    |                                 |
|                                                                                                                                                           |                    |                                 |
|                                                                                                                                                           |                    |                                 |
|                                                                                                                                                           |                    |                                 |
| Outcome Evaluation                                                                                                                                        |                    |                                 |
| Data Source/Method                                                                                                                                        | Sample Description | Expected Sample Size            |
|                                                                                                                                                           |                    |                                 |
|                                                                                                                                                           |                    |                                 |
|                                                                                                                                                           |                    |                                 |
|                                                                                                                                                           |                    |                                 |
| Section 4: Resources Needed                                                                                                                               |                    |                                 |
| A) Staffing Needs                                                                                                                                         |                    |                                 |
| <b>Evaluation activity</b><br>(e.g., develop recruitment and retention plan, design data collection forms/instruments, administer surveys)                |                    | <b>Staff member responsible</b> |
|                                                                                                                                                           |                    |                                 |
|                                                                                                                                                           |                    |                                 |
|                                                                                                                                                           |                    |                                 |
|                                                                                                                                                           |                    |                                 |
| B) Infrastructure Needs                                                                                                                                   |                    |                                 |
| <b>Data collection:</b> _____                                                                                                                             |                    |                                 |
| <b>Data storage:</b> _____                                                                                                                                |                    |                                 |
| <b>Data analysis:</b> _____                                                                                                                               |                    |                                 |
| <b>Reporting results:</b> _____                                                                                                                           |                    |                                 |
| C) Additional Costs                                                                                                                                       |                    |                                 |
| Identify additional costs associated with the evaluation beyond typical PSC implementation (e.g., additional staff, data storage, or management systems). |                    |                                 |
|                                                                                                                                                           |                    |                                 |
|                                                                                                                                                           |                    |                                 |
|                                                                                                                                                           |                    |                                 |
|                                                                                                                                                           |                    |                                 |

do not include all tasks that might be part of an evaluation, a representative set of tasks has been selected for each to demonstrate the use of this worksheet.

Now that you have started to design your evaluation, you should also consider putting together a more detailed timeline for the evaluation. A useful tool for establishing an evaluation timeline is a Gantt chart. In developing a Gantt chart for your evaluation, place the timeline at the top of the chart in the blue boxes and each evaluation activity in the column at the left. When creating the timeline, tailor the units of time to your evaluation—for instance, you could assign tasks by weeks or by months, depending on the duration of your evaluation. You can place an “X” or shade in the boxes to indicate when each evaluation activity will take

**Example L.I.N.K.S. Program Evaluation Plan**

| Section 1: Time Frame                                                                            |                                                                 |                             |                                     |                            |
|--------------------------------------------------------------------------------------------------|-----------------------------------------------------------------|-----------------------------|-------------------------------------|----------------------------|
| Expected start date: March 2017                                                                  |                                                                 |                             |                                     |                            |
| Expected end date: September 2018                                                                |                                                                 |                             |                                     |                            |
| Section 2: Methods/Data Sources                                                                  |                                                                 |                             |                                     |                            |
| Process Evaluation                                                                               |                                                                 |                             |                                     |                            |
| <b>Data Source/Method</b>                                                                        | Available now                                                   |                             | To be implemented                   |                            |
| 1. Attendance records (both in-person and based on registration website)                         | <input checked="" type="checkbox"/>                             |                             | <input type="checkbox"/>            |                            |
| 2. Personnel records (for staff qualifications)                                                  | <input type="checkbox"/>                                        |                             | <input checked="" type="checkbox"/> |                            |
| 3. Staff log of updates to curriculum                                                            | <input type="checkbox"/>                                        |                             | <input checked="" type="checkbox"/> |                            |
| 4. Satisfaction surveys                                                                          | <input type="checkbox"/>                                        |                             | <input checked="" type="checkbox"/> |                            |
| Data collection schedule:                                                                        |                                                                 |                             |                                     |                            |
| <b>Data Source/Method</b>                                                                        | Timing                                                          |                             |                                     |                            |
| 1. Attendance records                                                                            | Collect at end of every workshop                                |                             |                                     |                            |
| 2. Personnel records                                                                             | Monitor throughout evaluation period                            |                             |                                     |                            |
| 3. Staff log of updates to curriculum                                                            | Monitor throughout evaluation period                            |                             |                                     |                            |
| 4. Satisfaction surveys                                                                          | Administer at end of every workshop                             |                             |                                     |                            |
| Outcome Evaluation                                                                               |                                                                 |                             |                                     |                            |
| <b>Data Source/Method</b>                                                                        | Available now                                                   |                             | To be implemented                   |                            |
| 1. L.I.N.K.S. Challenge knowledge pretest and posttest                                           | <input checked="" type="checkbox"/>                             |                             | <input type="checkbox"/>            |                            |
| 2. Behavioral scales (relationship skills questionnaire, perceived efficacy to access resources) | <input type="checkbox"/>                                        |                             | <input checked="" type="checkbox"/> |                            |
| Data collection schedule:                                                                        |                                                                 |                             |                                     |                            |
| <b>Data Source/Method</b>                                                                        | <b>Pre-Participation</b>                                        | <b>During Participation</b> | <b>Post-Participation</b>           | <b>Follow-Up (specify)</b> |
| 1. L.I.N.K.S. Challenge test                                                                     | X                                                               |                             | X                                   |                            |
| 2. Relationship skills questionnaire                                                             | X                                                               |                             |                                     | X (6 months)               |
| 3. Self-efficacy to access resources                                                             | X                                                               |                             |                                     | X (6 months)               |
| Section 3: Selecting/Describing Sample                                                           |                                                                 |                             |                                     |                            |
| Process Evaluation                                                                               |                                                                 |                             |                                     |                            |
| <b>Data Source/Method</b>                                                                        | <b>Sample Description</b>                                       |                             | <b>Expected Sample Size</b>         |                            |
| 1. Attendance records                                                                            | All participants completing workshop from July to December 2017 |                             | ~100 individuals                    |                            |
| 2. Personnel records                                                                             | For all L.I.N.K.S. personnel required to have credentials       |                             | 10 staff members                    |                            |

|                                                                                                                                                                                                                                                                                                     |                                                                 |                      |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------|----------------------|
| 3. Log of curriculum updates                                                                                                                                                                                                                                                                        | Log will track all curriculum updates during evaluation period  | N/A                  |
| 4. Satisfaction surveys                                                                                                                                                                                                                                                                             | All participants completing workshop from July to December 2017 | ~100 individuals     |
| Outcome Evaluation                                                                                                                                                                                                                                                                                  |                                                                 |                      |
| Data Source/Method                                                                                                                                                                                                                                                                                  | Sample Description                                              | Expected Sample Size |
| 1. L.I.N.K.S. Challenge                                                                                                                                                                                                                                                                             | All participants completing workshop from July to December 2017 | ~100                 |
| 2. Behavioral scales (relationship skills, self-efficacy)                                                                                                                                                                                                                                           | All participants completing workshop from July to December 2017 | ~100                 |
| Section 4: Resources Needed                                                                                                                                                                                                                                                                         |                                                                 |                      |
| A) Staffing Needs                                                                                                                                                                                                                                                                                   |                                                                 |                      |
| Evaluation activity                                                                                                                                                                                                                                                                                 | Staff member responsible                                        |                      |
| 1. Administer pre- and post-surveys                                                                                                                                                                                                                                                                 | 1. Workshop leader (depends on workshop)                        |                      |
| 2. Program behavioral scales for follow-up into online survey, send link to participants                                                                                                                                                                                                            | 2. [name here]                                                  |                      |
| 3. Review personnel records for credentials, monitor updates to curriculum                                                                                                                                                                                                                          | 3. [name here]                                                  |                      |
| 4. Enter data from surveys into spreadsheet                                                                                                                                                                                                                                                         | 4. [name here]                                                  |                      |
| 5. Analyze and report on data to leadership                                                                                                                                                                                                                                                         | 5. [name here]                                                  |                      |
| B) Infrastructure Needs                                                                                                                                                                                                                                                                             |                                                                 |                      |
| <b>Data collection:</b> Photocopied knowledge tests for pre- and post-surveys; log to track updates to curriculum (create spreadsheet on staff computer); online survey system for follow-up measures + access to participant emails to follow-up. Possible to get from online registration system. |                                                                 |                      |
| <b>Data storage:</b> Develop spreadsheet system to enter data for knowledge tests and behavioral scales to auto-score.                                                                                                                                                                              |                                                                 |                      |
| <b>Data analysis:</b> Data analysis can be conducted in Excel with RAND tool.                                                                                                                                                                                                                       |                                                                 |                      |
| <b>Reporting results:</b> Can use existing software to summarize results at end of evaluation period. Results will be reported internally and to leadership per their interest.                                                                                                                     |                                                                 |                      |
| C) Additional Costs                                                                                                                                                                                                                                                                                 |                                                                 |                      |
| <b>Identify additional costs associated with the evaluation beyond typical PSC implementation (e.g., additional staff, data storage, or management systems).</b>                                                                                                                                    |                                                                 |                      |
| Will use SurveyMonkey for follow-up behavioral scale, so no cost involved. Staff time (approximately two weeks) needed for analysis and reporting of results.                                                                                                                                       |                                                                 |                      |

**Example FMEAP Service Evaluation Plan**

| Section 1: Time Frame                                                                                      |                                                                                                                                                                                                                          |                             |                                     |                            |
|------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|-------------------------------------|----------------------------|
| <b>Expected start date:</b> November 2016                                                                  |                                                                                                                                                                                                                          |                             |                                     |                            |
| <b>Expected end date:</b> April 2018                                                                       |                                                                                                                                                                                                                          |                             |                                     |                            |
| Section 2: Methods/Data Sources                                                                            |                                                                                                                                                                                                                          |                             |                                     |                            |
| Process Evaluation                                                                                         |                                                                                                                                                                                                                          |                             |                                     |                            |
| <b>Data Source/Method</b>                                                                                  | Available now                                                                                                                                                                                                            |                             | To be implemented                   |                            |
| 1. Sign-in sheets for training sessions                                                                    | <input checked="" type="checkbox"/>                                                                                                                                                                                      |                             | <input type="checkbox"/>            |                            |
| 2. Appointment calendars of career coaches                                                                 | <input checked="" type="checkbox"/>                                                                                                                                                                                      |                             | <input type="checkbox"/>            |                            |
| 3. Log of walk-ins requesting referrals                                                                    | <input type="checkbox"/>                                                                                                                                                                                                 |                             | <input checked="" type="checkbox"/> |                            |
| 4. Log of telephone calls requesting referrals                                                             | <input type="checkbox"/>                                                                                                                                                                                                 |                             | <input checked="" type="checkbox"/> |                            |
| 5. Personnel records for record of continuing education                                                    | <input checked="" type="checkbox"/>                                                                                                                                                                                      |                             | <input type="checkbox"/>            |                            |
| 6. Satisfaction surveys                                                                                    | <input type="checkbox"/>                                                                                                                                                                                                 |                             | <input checked="" type="checkbox"/> |                            |
| <b>Data collection schedule:</b>                                                                           |                                                                                                                                                                                                                          |                             |                                     |                            |
| <b>Data Source/Method</b>                                                                                  | <b>Timing</b>                                                                                                                                                                                                            |                             |                                     |                            |
| 1. Sign-in sheets for training sessions                                                                    | Collect at end of every training session                                                                                                                                                                                 |                             |                                     |                            |
| 2. Appointment calendars of career coaches                                                                 | Monitor on monthly basis                                                                                                                                                                                                 |                             |                                     |                            |
| 3. Log of walk-ins requesting referrals                                                                    | Collate on monthly basis                                                                                                                                                                                                 |                             |                                     |                            |
| 4. Log of telephone calls requesting referrals                                                             | Collate on monthly basis                                                                                                                                                                                                 |                             |                                     |                            |
| 5. Personnel records for record of continuing education                                                    | Monitor throughout evaluation                                                                                                                                                                                            |                             |                                     |                            |
| 6. Satisfaction surveys                                                                                    | Administer at end of training sessions and coaching sessions. For those requesting referral alone, provide survey at end of session when in-person. Over phone, program one-question satisfaction survey at end of call. |                             |                                     |                            |
| Outcome Evaluation                                                                                         |                                                                                                                                                                                                                          |                             |                                     |                            |
| <b>Data Source/Method</b>                                                                                  | Available now                                                                                                                                                                                                            |                             | To be implemented                   |                            |
| 1. Knowledge tests for career training sessions                                                            | <input checked="" type="checkbox"/>                                                                                                                                                                                      |                             | <input type="checkbox"/>            |                            |
| 2. Checklist of accomplishments at end of career coaching (e.g., completed resume, submitted applications) | <input type="checkbox"/>                                                                                                                                                                                                 |                             | <input checked="" type="checkbox"/> |                            |
| 3. Follow-up surveys to assess career outcomes                                                             | <input type="checkbox"/>                                                                                                                                                                                                 |                             | <input checked="" type="checkbox"/> |                            |
| <b>Data collection schedule:</b>                                                                           |                                                                                                                                                                                                                          |                             |                                     |                            |
| <b>Data Source/Method</b>                                                                                  | <b>Pre-Participation</b>                                                                                                                                                                                                 | <b>During Participation</b> | <b>Post-Participation</b>           | <b>Follow-Up (specify)</b> |
| 1. Knowledge tests                                                                                         | X                                                                                                                                                                                                                        |                             | X                                   |                            |
| 2. Checklist of accomplishments                                                                            |                                                                                                                                                                                                                          |                             | X                                   |                            |
| 3. Follow-up surveys to assess career outcomes                                                             |                                                                                                                                                                                                                          |                             |                                     | X (6 months)               |

| Section 3: Selecting/Describing Sample                                                                  |                                                                                 |                                                                      |
|---------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------|----------------------------------------------------------------------|
| Process Evaluation                                                                                      |                                                                                 |                                                                      |
| Data Source/Method                                                                                      | Sample Description                                                              | Expected Sample Size                                                 |
| 1. Sign-in sheets for training sessions                                                                 | All participants completing workshop from January to April 2017                 | ~60 family members                                                   |
| 2. Appointment calendars of career coaches                                                              | For all coaches providing coaching sessions from January to April 2017          | 8 career coaches, expected to result in 30 individuals using service |
| 3. Log of walk-ins requesting referrals                                                                 | Monitor from January to April 2017                                              | ~50 individuals                                                      |
| 4. Log of telephone calls requesting referrals                                                          | Monitor from January to April 2017                                              | ~40 individuals                                                      |
| 5. Personnel records for record of continuing education                                                 | For all FMEAP personnel required to participate in annual continuing education  | 10 staff members                                                     |
| 6. Satisfaction surveys                                                                                 | All individuals using FMEAP services from January to April 2017                 | ~180 individuals                                                     |
| Outcome Evaluation                                                                                      |                                                                                 |                                                                      |
| Data Source/Method                                                                                      | Sample Description                                                              | Expected Sample Size                                                 |
| 1. Knowledge tests                                                                                      | All participants completing training sessions from January to April 2017        | ~60                                                                  |
| 2. Checklist of accomplishments                                                                         | All participants completing career coaching sessions from January to April 2017 | ~30                                                                  |
| 3. Follow-up surveys to assess career outcomes                                                          | All individuals using FMEAP services from January to April 2017                 | ~180 individuals                                                     |
| Section 4: Resources Needed                                                                             |                                                                                 |                                                                      |
| A) Staffing Needs                                                                                       |                                                                                 |                                                                      |
| Evaluation activity                                                                                     | Staff member responsible                                                        |                                                                      |
| 1. Track sign-in sheets and obtain contact information for participants                                 | 1. Workshop leader (depends on workshop)                                        |                                                                      |
| 2. Provide data about career coaching sessions                                                          | 2. All career coaches                                                           |                                                                      |
| 3. Track that all staff have participated in continuing education                                       | 3. [name here]                                                                  |                                                                      |
| 4. Log all phone calls for referrals                                                                    | 4. All staff answering telephones                                               |                                                                      |
| 5. Aggregate referral data from walk-ins and phone                                                      | 5. [name here]                                                                  |                                                                      |
| 6. Develop satisfaction survey and determine way to administer over phone to those calling for referral | 6. FMEAP manager                                                                |                                                                      |
| 7. Develop knowledge tests                                                                              | 7. Workshop leaders (based on curriculum)                                       |                                                                      |
| 8. Administer satisfaction surveys and knowledge tests                                                  | 8. Workshop leaders and career coaches                                          |                                                                      |
| 9. Develop follow-up survey                                                                             | 9. [name here]                                                                  |                                                                      |
| 10. Maintain list of participant contact information and send follow-up survey via email and/or mail    | 10. [name here]                                                                 |                                                                      |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| B) Infrastructure Needs                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| <p><b>Data collection:</b> Need to develop system to track phone calls requesting referrals, and method to deliver brief satisfaction survey via phone to these individuals and automate the collection responses. Need to develop instrument corresponding to education/training sessions to assess knowledge. Need to maintain list of participant contact information, including preferred method (email or mail).</p> <p><b>Data storage:</b> Develop spreadsheet system to enter data for knowledge test and satisfaction survey. Develop system to track individual goals and long-term outcomes for individuals attending career-coaching sessions (i.e., so that it is possible to tell if someone was submitting job applications, what the success rate was for interviews and/or job offers).</p> <p><b>Data analysis:</b> Data analysis can be conducted in Excel with RAND tool. Need staff member familiar enough with Excel to conduct these analyses.</p> <p><b>Reporting results:</b> Results will be shared with Personal &amp; Professional Development leadership at base. Need staff member to create briefing.</p> |
| C) Additional Costs                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| <p><b>Identify additional costs associated with the evaluation beyond typical PSC implementation (e.g., additional staff, data storage, or management systems).</b></p> <p>Mailing costs for participants who prefer to receive follow-up communication via mail instead of email. Cost for telephone satisfaction survey. Staff time for tracking participants, particularly those receiving referral services alone because this is a large subset of FMEAP participants, and there are not already systems in place to collect identifiable information from these individuals that would be needed for any follow-up communications.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |

**Example Outdoor Recreation Equipment Checkout and Equipment Rental Capability Evaluation Plan**

| Section 1: Time Frame                                                  |                                         |                             |                                     |                            |
|------------------------------------------------------------------------|-----------------------------------------|-----------------------------|-------------------------------------|----------------------------|
| <b>Expected start date:</b> July 2016                                  |                                         |                             |                                     |                            |
| <b>Expected end date:</b> February 2019                                |                                         |                             |                                     |                            |
| Section 2: Methods/Data Sources                                        |                                         |                             |                                     |                            |
| Process Evaluation                                                     |                                         |                             |                                     |                            |
| <b>Data Source/Method</b>                                              | <b>Available now</b>                    |                             | <b>To be implemented</b>            |                            |
| 1. Registration forms for equipment                                    | <input checked="" type="checkbox"/>     |                             | <input type="checkbox"/>            |                            |
| 2. Invoices for rentals                                                | <input checked="" type="checkbox"/>     |                             | <input type="checkbox"/>            |                            |
| 3. Logs of participants in safety and skills lessons                   | <input type="checkbox"/>                |                             | <input checked="" type="checkbox"/> |                            |
| 4. Personnel records (for staff qualifications)                        | <input checked="" type="checkbox"/>     |                             | <input type="checkbox"/>            |                            |
| 5. Equipment safety inventory records                                  | <input checked="" type="checkbox"/>     |                             | <input checked="" type="checkbox"/> |                            |
| 6. Satisfaction surveys                                                | <input type="checkbox"/>                |                             | <input checked="" type="checkbox"/> |                            |
| <b>Data collection schedule:</b>                                       |                                         |                             |                                     |                            |
| <b>Data Source/Method</b>                                              | <b>Timing</b>                           |                             |                                     |                            |
| 1. Registration forms for equipment                                    | Collate every month                     |                             |                                     |                            |
| 2. Invoices for rentals                                                | Collate every month                     |                             |                                     |                            |
| 3. Logs of participants in safety and skills lessons                   | Collate every month                     |                             |                                     |                            |
| 4. Personnel records (for staff qualifications)                        | Monitor throughout evaluation           |                             |                                     |                            |
| 5. Equipment safety inventory records                                  | Monitor throughout evaluation           |                             |                                     |                            |
| 6. Satisfaction surveys                                                | Administer after rental has taken place |                             |                                     |                            |
| <b>Outcome Evaluation</b>                                              |                                         |                             |                                     |                            |
| <b>Data Source/Method</b>                                              | <b>Available now</b>                    |                             | <b>To be implemented</b>            |                            |
| 1. Invoices for rentals                                                | <input checked="" type="checkbox"/>     |                             | <input type="checkbox"/>            |                            |
| 2. Quiz to measure safety and equipment skills                         | <input type="checkbox"/>                |                             | <input checked="" type="checkbox"/> |                            |
| 3. Follow-up survey to assess quality of life, community connectedness | <input type="checkbox"/>                |                             | <input checked="" type="checkbox"/> |                            |
| <b>Data collection schedule:</b>                                       |                                         |                             |                                     |                            |
| <b>Data Source/Method</b>                                              | <b>Pre-Participation</b>                | <b>During Participation</b> | <b>Post-Participation</b>           | <b>Follow-Up (specify)</b> |
| 1. Invoices for rentals                                                |                                         |                             | X                                   |                            |
| 2. Quiz to measure safety and equipment skills                         |                                         |                             | X                                   |                            |
| 3. Follow-up survey to assess quality of life, community connectedness |                                         |                             |                                     | X (6 months)               |

| Section 3: Selecting/Describing Sample                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                   |                                                                                                                                                                                                                                                                                                               |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Process Evaluation                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                   |                                                                                                                                                                                                                                                                                                               |
| Data Source/Method                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Sample Description                                                                | Expected Sample Size                                                                                                                                                                                                                                                                                          |
| 1. Registration forms for equipment                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | All individuals renting equipment from April to August 2017                       | ~400 individuals                                                                                                                                                                                                                                                                                              |
| 2. Invoices for rentals                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | All invoices from equipment rentals from April to August 2017                     | ~320 invoices                                                                                                                                                                                                                                                                                                 |
| 3. Logs of participants in safety and skills lessons                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | Monitor from April to August 2017                                                 | ~100 individuals                                                                                                                                                                                                                                                                                              |
| 4. Personnel records (for staff qualifications)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | Monitor throughout evaluation period                                              | 12 staff members                                                                                                                                                                                                                                                                                              |
| 5. Equipment safety inventory records                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Monitor throughout evaluation period                                              | Safety inventories taken once every month                                                                                                                                                                                                                                                                     |
| 6. Satisfaction surveys                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | All individuals using capability from April to August 2017                        | ~275 individuals (may be challenging to obtain satisfaction surveys from all participants)                                                                                                                                                                                                                    |
| Outcome Evaluation                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                   |                                                                                                                                                                                                                                                                                                               |
| Data Source/Method                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Sample Description                                                                | Expected Sample Size                                                                                                                                                                                                                                                                                          |
| 1. Invoices for rentals                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | All individuals renting equipment from April to August 2017                       | ~400 individuals                                                                                                                                                                                                                                                                                              |
| 2. Quiz to measure safety and equipment skills                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | All individuals taking part in skills and safety lesson from April to August 2017 | ~100 individuals                                                                                                                                                                                                                                                                                              |
| 3. Follow-up survey to assess quality of life, community connectedness                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | All individuals using equipment rental services from January to April 2017        | ~150–200 individuals (expect somewhat low response rate, since many use this service for a single rental and may not be interested in taking survey)                                                                                                                                                          |
| Section 4: Resources Needed                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                                                                                   |                                                                                                                                                                                                                                                                                                               |
| A) Staffing Needs                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                   |                                                                                                                                                                                                                                                                                                               |
| <b>Evaluation activity</b><br>1. Modify registration form to ensure that relevant demographic data are collected (e.g., who is renting equipment—Marine vs. command vs. dependent rental)<br>2. Enter data from registration forms into electronic form to track types of equipment rentals<br>3. Ensure that safety inventories are being conducted and recorded periodically<br>4. Design brief knowledge quiz to accompany skills and safety lessons<br>5. Design one-question satisfaction survey to be handed to participants when they return rental<br>6. Design brief follow-up survey to be sent to participants<br>7. Create system to track preferred contact information for distribution of follow-up survey<br>8. Identify potential comparison group for assessing community connectedness |                                                                                   | <b>Staff member responsible</b><br>1. Equipment Rental Manager<br>2. All rental staff interacting with customers<br>3. Equipment Rental Manager<br>4. Equipment Rental Manager<br>5. Equipment Rental Manager<br>6. [name here]<br>7. [name here]<br>8. Coordinate with Semper Fit leadership at installation |
| B) Infrastructure Needs                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                   |                                                                                                                                                                                                                                                                                                               |
| <b>Data collection:</b> Need to develop system to contact participants after using the service. Would be ideal to contact via email and send a link to a brief survey.<br><b>Data storage:</b> System needed to electronically store registration form info. Could be simple spreadsheet. Would be helpful to automate safety inventory results too.<br><b>Data analysis:</b> Data analysis can be conducted in Excel with RAND tool.<br><b>Reporting results:</b> No additional infrastructure needs.                                                                                                                                                                                                                                                                                                    |                                                                                   |                                                                                                                                                                                                                                                                                                               |

## C) Additional Costs

**Identify additional costs associated with the evaluation beyond typical PSC implementation (e.g., additional staff, data storage, or management systems).**

Staff time for sending follow-up surveys, particularly if it is not possible to automate via email. Need to identify appropriate comparison group of individuals not using Equipment Rentals to compare community connectedness. Semper Fit leadership may be able to help us figure out how to access this group. Will be additional logistical needs associated with reaching out to those individuals and, potentially, costs as a result.

place. You can use Worksheet 4.4 to create your own Gantt chart. A sample Gantt chart for the L.I.N.K.S. evaluation plan is provided after Worksheet 4.4.

## Summary

This chapter provided guidance about the purpose of process and outcome evaluations, as well as design elements related to both types of evaluation. We also discussed the types of data that you may use for an evaluation. We reviewed some of the considerations that guide evaluation design, including treatment of human subjects,<sup>5</sup> and you had the opportunity to think about how these apply to your program (Worksheet 4.2). Finally, you began to identify the key elements of your evaluation (Worksheet 4.3) and create a timeline for your evaluation (Worksheet 4.4). Chapter Five will help you to get more specific about how you will define your process and outcome measures.

<sup>5</sup> Federal regulations 45 Code of Federal Regulations 46.102 (Protection of Human Subjects 2009) and Secretary of the Navy Instruction 3900.39D define research as “a systematic investigation including research development, testing, and evaluation, designed to develop or contribute to generalizable knowledge.” Human subject means a living individual about whom an investigator (whether professional or student) conducting research obtains (1) data through intervention or interaction with the individual, or (2) identifiable private information. Program evaluations are considered aspects of research and *must* be vetted through the HRPP *before* beginning. The program, staff, or investigators themselves cannot make study determinations and must use the HRPP Applicability Review Process for all potential research and program evaluations.

Worksheet 4.4  
Evaluation Gantt Chart

|      | Timeline |  |  |  |  |  |  |  |  |  |  |  |
|------|----------|--|--|--|--|--|--|--|--|--|--|--|
| Task |          |  |  |  |  |  |  |  |  |  |  |  |
|      |          |  |  |  |  |  |  |  |  |  |  |  |
|      |          |  |  |  |  |  |  |  |  |  |  |  |
|      |          |  |  |  |  |  |  |  |  |  |  |  |
|      |          |  |  |  |  |  |  |  |  |  |  |  |
|      |          |  |  |  |  |  |  |  |  |  |  |  |
|      |          |  |  |  |  |  |  |  |  |  |  |  |
|      |          |  |  |  |  |  |  |  |  |  |  |  |
|      |          |  |  |  |  |  |  |  |  |  |  |  |
|      |          |  |  |  |  |  |  |  |  |  |  |  |
|      |          |  |  |  |  |  |  |  |  |  |  |  |
|      |          |  |  |  |  |  |  |  |  |  |  |  |
|      |          |  |  |  |  |  |  |  |  |  |  |  |

Example GANTT Chart for L.I.N.K.S. Program

|                                                                                                   | Timeline (month/year) |      |      |      |      |      |      |       |       |       |      |      |      |      |      |      |      |      |      |
|---------------------------------------------------------------------------------------------------|-----------------------|------|------|------|------|------|------|-------|-------|-------|------|------|------|------|------|------|------|------|------|
| Task                                                                                              | 3/17                  | 4/17 | 5/17 | 6/17 | 7/17 | 8/17 | 9/17 | 10/17 | 11/17 | 12/17 | 1/18 | 2/18 | 3/18 | 4/18 | 5/18 | 6/18 | 7/18 | 8/18 | 9/18 |
| 1. Develop spreadsheet to enter data for knowledge tests and behavioral scales                    | X                     | X    |      |      |      |      |      |       |       |       |      |      |      |      |      |      |      |      |      |
| 2. Photocopy L.I.N.K.S. Challenge pretest/posttest                                                |                       |      | X    | X    | X    | X    | X    | X     | X     | X     |      |      |      |      |      |      |      |      |      |
| 3. Program online survey for behavioral scales (for 6-month follow-up)                            |                       |      |      |      | X    | X    |      |       |       |       |      |      |      |      |      |      |      |      |      |
| 4. Create log to track curriculum updates                                                         | X                     | X    |      |      |      |      |      |       |       |       |      |      |      |      |      |      |      |      |      |
| 5. Create system to track participant email address to be merged into online survey (using Excel) |                       | X    | X    |      |      |      |      |       |       |       |      |      |      |      |      |      |      |      |      |
| 6. Collect data from participants immediately before, during, and after workshop                  |                       |      |      | X    | X    | X    | X    | X     | X     | X     |      |      |      |      |      |      |      |      |      |
| 7. Collect follow-up data                                                                         |                       |      |      |      |      |      |      |       |       | X     | X    | X    | X    | X    | X    | X    |      |      |      |
| 8. Review personnel logs to ensure staff are credentialed                                         |                       |      |      | X    | X    | X    | X    | X     | X     | X     |      |      |      |      |      |      |      |      |      |
| 9. Monitor curriculum updates                                                                     |                       |      |      | X    | X    | X    | X    | X     | X     | X     |      |      |      |      |      |      |      |      |      |
| 10. Analyze data                                                                                  |                       |      |      |      |      |      |      |       |       |       | X    | X    | X    | X    | X    | X    | X    | X    |      |
| 11. Report results                                                                                |                       |      |      |      |      |      |      |       |       |       |      |      |      |      |      |      |      | X    | X    |



## Evaluation Measures

---

An important aspect of designing an evaluation is the selection of evaluation measures. In this section, we discuss two types of evaluation measures: measures of performance, used to evaluate program usage and implementation; and measures of effectiveness, designed to measure short-term and intermediate outcomes. In this chapter, we will discuss each type of measure. We will then review considerations for selecting measures and provide you with a template to identify measures appropriate for your evaluation.

### Measures of Performance and Effectiveness

You should now be familiar with the differences between process evaluation and outcome evaluation. Measures of performance (MOPs) and effectiveness (MOEs) correspond to each type of evaluation.

MOPs are a component of process evaluations and are used to evaluate the usage and implementation of a program. On the logic model, this corresponds to the program inputs, activities, and outputs (see Figure 5.1). For example, MOPs might address questions related to what program activities are implemented, how many people participated in the activities, and how satisfied program participants are. Because MOPs are focused on the project inputs, activities, and outputs, they have a more immediate time frame, which means that you will likely be focused on collecting data as the PSC is being implemented. For this reason, MOPs can also be analyzed at earlier phases of your evaluation.

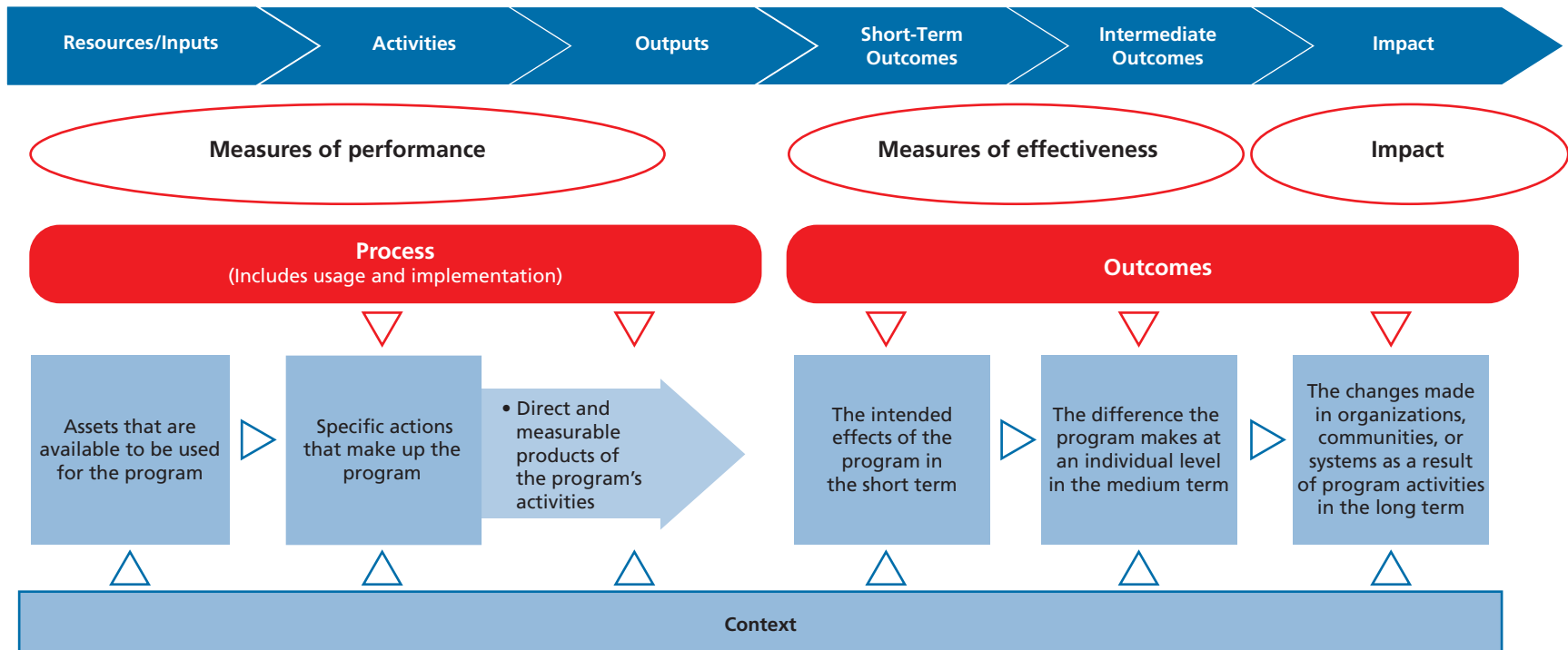
MOEs are a component of outcome evaluations and are used to measure whether a PSC has achieved the intended outcomes. MOEs allow organizations to determine what participant-level changes resulted from PSC activities. These outcomes may be observed in the short term but may also be observed over an intermediate time frame.

A summary of the differences between MOPs and MOEs is presented in Table 5.1. For MF PSCs, both MOPs and MOEs will generally be measured at the individual or family level.

### Defining Measures of Performance

Your logic model can be used as a tool to help identify MOPs. Recall that process evaluation focuses on the input, activity, and output portion of the logic model. Therefore, to identify MOPs, you should begin by revisiting your logic model. You have already identified key elements of the implementation, such as the target population and core components of the

**Figure 5.1**  
**Connection Between Logic Model and Evaluation Measures**



RAND TL264-5.1

**Table 5.1**  
**Characteristics of Measures of Performance and Measures of Effectiveness**

|                     | Measures of Performance (MOPs)                                                                              | Measures of Effectiveness (MOEs)                                 |
|---------------------|-------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------|
| Type of evaluation  | Process/implementation evaluation                                                                           | Outcome/impact evaluation                                        |
| Questions addressed | What activities are implemented? How many people participate in activities? How satisfied are participants? | What participant-level changes resulted from program activities? |
| Level of analysis   | Individual or family level                                                                                  | Individual or family level                                       |
| Time frame          | Immediate                                                                                                   | Short term and intermediate                                      |

**Table 5.2**  
**Sample MOPs for Evaluating Program Participation**

| Offering                                           | Measure of Performance                                                                                                                                                                                                                                                                          |
|----------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| L.I.N.K.S. Program                                 | <ol style="list-style-type: none"> <li>1. Number of individuals who attended a session, by session</li> <li>2. Proportion of interested families who were placed into a session</li> <li>3. Proportion of available spaces in a given session that were filled</li> </ol>                       |
| FMEAP Service                                      | <ol style="list-style-type: none"> <li>1. Number of individuals attending career training or education sessions</li> <li>2. Average number of sessions attended per individual</li> <li>3. Proportion of individuals interested in career training or education who attend a session</li> </ol> |
| Outdoor Recreation and Equipment Rental Capability | <ol style="list-style-type: none"> <li>1. Number of equipment checkouts, by type of equipment</li> <li>2. Number of individuals receiving skills and safety lessons</li> <li>3. Proportion of individuals required to receive skills and safety lessons who did receive lessons</li> </ol>      |

program (e.g., referrals provided, services provided, implementation of workshops, equipment rental made available). These will dictate your MOPs.

The next step is to figure out how each element should be defined. For example, you may want to measure participation in your PSC—but what does that mean? It is important to acknowledge that there are multiple ways that an MOP may be operationalized. For example, if we were interested in measuring the use of the L.I.N.K.S. program, there are many ways we could define that MOP. We might be interested in (a) the total number of people who used the program, (b) the number of people from the target population who used the program, and/or (c) the proportion of available slots in each workshop that were filled. In Table 5.2, we provide examples of different ways you could measure *utilization* for each of our sample offerings. This is why it is so important to be specific when you develop your logic model. The clearer you are in your logic model, the easier it will be to identify and define your MOPs.

It is also important to keep in mind that many offerings in the MF portfolio have multiple components. Outdoor Recreation and Equipment Rental Capability is a good example of this. Not only is this service designed to provide equipment rentals, but it also provides skills

and safety lessons. In this case, different MOPs would be needed to measure each aspect of the capability.

### Measure Considerations

When selecting MOPs and thinking about how you will operationalize those MOPs, there are a number of other important considerations. In this section, we review four considerations that you will want to keep in mind as you identify your MOPs.

1. **Identifying and selecting a data source.** We reviewed the distinction between primary and secondary data in Chapter Four. As you identify your MOPs, be sure to consider the sources of data available. If you are measuring attendance, are there attendance logs available, and are they consistent across events and locations? Is there an online registration portal? If you are interested in measuring fidelity to a program's curriculum, is there someone familiar with the program who could observe sessions and rate fidelity? If you are measuring satisfaction, is there a way to administer a satisfaction survey to participants or people who use a given service or capability? How do you reduce participant burden while still collecting the data you need? These types of considerations will help you to operationalize your MOPs.
2. **Applying the SMART framework.** We discussed the importance of the SMART principles when developing your logic model in Chapter Three. Using the SMART principles to describe your program generally makes it easier to define your MOPs as well. Consider FMEAP as an example. Perhaps you are interested in measuring whether the PSC reached its target audience, which you define as the proportion of the target population who contacted the program. To obtain a meaningful MOP, it will be key to have a specific definition of who the target audience is. If you simply indicated that the target population is Marine family members and defined your measure as the number of individuals who contacted the program divided by the number of Marine family members, then your measure would not be completely accurate: It would include individuals who were in need of employment assistance services and those who were not, and you may conclude that FMEAP is reaching a smaller proportion of the target population than it is. If, instead, you more narrowly define the target population as Marine family members seeking employment assistance services, and your measure is the number of people who contact FMEAP divided by the number of *interested* Marine family members, then you will have more meaningful information about the service.
3. **Identifying and establishing metrics.** Some PSCs have specific requirements related to their implementation. For example, an MCO may specify that a workshop must be offered a certain number of times in a year, or state that all Marines in a certain role should participate in a program within a certain time frame. These specific targets are called *metrics*. For example, staff members of the FMEAP program are required to have appropriate professional certifications. Therefore, one MOP might be to calculate the proportion of FMEAP staff required to have professional certifications that actually do have the necessary credentials. For this measure, an appropriate metric would be 100 percent, as all staff should have these required qualifications. Metrics are useful because they provide a standard against which to compare the performance of your PSC.

Not all of the MOPs that you identify will have a specific target. It may be that the goal is to reach as many family members as possible, or to increase the capacity for services from year to year, but with no specific target attached to the measure. In this case, you may choose to simply track the measure over time.

4. **Determining the time frame for data analysis.** It is also important to consider the time frame over which data should be analyzed. In part, this may be determined by the metrics you have identified. For example, if your program is required to offer a workshop a certain number of times per year, it may be appropriate to examine MOPs on an annual basis. However, not all PSCs have specific guidance regarding the number of participants reached or the number of activities provided in a certain time frame. Therefore, you should consider what type of time frame makes the most sense for your own PSC. Some questions that may help guide this decision:
  - **How often are PSC activities taking place?** If your PSC is constantly providing services, then evaluating the implementation over a shorter period of time (e.g., every three months) may be appropriate. Alternatively, if your PSC offers workshops that take place only twice each year, collecting data annually may be appropriate.
  - **What resources are available for data collection and analysis?** If software systems are available for data collection, it may be possible to collect and analyze data in real time (e.g., if electronic sign-in systems are used for workshops). However, if data are collected manually and then compiled, it may not be possible to collect and analyze data as frequently.
  - **How many people generally use the program?** If your program serves large numbers of individuals, it may be possible to examine the implementation over a shorter time frame. However, if only a small number of people use the program, it may make more sense to track implementation over a longer time frame so that it is easier to identify trends.

### Sample MOP Definitions

In Tables 5.3, 5.4, and 5.5, we identify MOPs for the three sample PSCs. In the first column, we identify the MOPs that correspond to the logic models presented in Chapter Three. In the second column, we provide a definition for each MOP. These are the specific ways that you could operationalize each MOP. You will notice that many of the definitions that we provide are in the format of a numerator and denominator. For example, for L.I.N.K.S., the second MOP is whether the sessions meet demand. We define this as the number of families placed into a session divided by the number of eligible families who contact L.I.N.K.S. to register for any session. More simply put, this is the proportion of interested and eligible families placed into a session. However, by specifically indicating the numerator and denominator, the method of calculating this proportion is clearer.

### Identifying MOPs for Your Program

You have already identified the inputs, activities, and outputs from your PSC, which is the first step to identifying MOPs. Next, using information from your logic model, fill in Worksheet 5.1. We have provided examples based on our three sample PSCs and the MOP lists above.

In the first column, Input/Activity/Output, you will fill in the elements from the logic model you created in Chapter Three. In the second column, identify your MOP. In the third

column, identify the way that you will define each MOP. In the fourth column, you can identify a specific metric if one is indicated for the MOP. If there is not a specific metric, you can include details about the time frame over which the MOP should be measured and/or analyzed. Finally, in the fifth column, indicate the data source. Remember to consider both primary and secondary source of data. If there is not currently a data source that can be used for the MOP, consider the ways that your program could begin to collect the data. After you complete this worksheet, you may consider returning to Worksheet 4.3 to add more detail about the measures and data sources that you will use for your process evaluation.

**Table 5.3**  
**L.I.N.K.S. Program MOPs**

| Measure of Performance                                       | Definition                                                                                                                     |
|--------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|
| Inputs/Resources                                             |                                                                                                                                |
| Staff are adequately trained                                 | Number of staff with required level of training/number of staff<br><i>Note: The metric is 100%</i>                             |
| Frequency of session content updates                         | Number of times that substantive content is added, removed, or revised per year (excluding minor edits to text or images)      |
| Activities/Outputs                                           |                                                                                                                                |
| Number of sessions offered per year                          | Number of sessions taking place each year, for each type of session (e.g., 2-day spouse sessions, 3-hour child sessions)       |
| Proportion of interested families who attended a session     | Number of families attending a session/Number of eligible families who contact L.I.N.K.S. to register for any session          |
| Average attendance, by type of session                       | Number of families who attend a session/Number of sessions, calculated by session type                                         |
| Number of families who attend any type of L.I.N.K.S. session | Number of unique families who attend any type of L.I.N.K.S. session, by year                                                   |
| Participant ratings of satisfaction                          | Participant ratings of (1) liking program, (2) finding program to be personally relevant, and (3) finding program to be useful |

**Table 5.4**  
**FMEAP Service MOPs**

| Measure of Performance                                                                                                                                | Definition                                                                                                                                                                                                                                                                                         |
|-------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Inputs/Resources</b>                                                                                                                               |                                                                                                                                                                                                                                                                                                    |
| Staff have appropriate qualifications                                                                                                                 | Number of FMEAP staff with professional certifications/<br>Number of FMEAP staff required to have professional certification<br><i>Note: The metric is 100%</i>                                                                                                                                    |
| Staff participate in continuing education                                                                                                             | 1. Number of staff participating in continuing education each year/Total number of staff<br>2. Number of continuing education programs attended by staff per year                                                                                                                                  |
| Availability of relevant resources (e.g., books, publications, periodicals, videos/online programs)                                                   | Number of resources available by type                                                                                                                                                                                                                                                              |
| <b>Activities/Outputs</b>                                                                                                                             |                                                                                                                                                                                                                                                                                                    |
| Frequency of outreach to local businesses, professional organizations, and employment offices                                                         | Number of contacts with local businesses, professional organizations, or employment offices                                                                                                                                                                                                        |
| Number of job fairs FMEAP hosts or participates in                                                                                                    | Number of career/job fairs hosted or participated in annually<br><i>Note: The metric is at least twice annually.</i>                                                                                                                                                                               |
| Number of career training or education sessions                                                                                                       | Total number of sessions (overall and by topic)                                                                                                                                                                                                                                                    |
| Number of family members attending career training or education sessions                                                                              | 1. Number of individuals attending career training or education sessions<br>2. Average number of sessions per individual                                                                                                                                                                           |
| Proportion of interested family members attending career training or education sessions                                                               | Number of individuals attending career training or education/Number of individuals interested in career training or education                                                                                                                                                                      |
| Number of family members receiving career coaching sessions                                                                                           | 1. Number of family members receiving career coaching sessions<br>2. Average number of sessions per individual by modality (i.e., group vs. one-on-one)                                                                                                                                            |
| Proportion of interested family members attending career coaching sessions                                                                            | Number of individuals receiving career coaching/<br>Number of individuals interested in career coaching                                                                                                                                                                                            |
| Proportion of interested family members receiving referrals to additional resources or opportunities (e.g., volunteer opportunities, career websites) | Number of individuals receiving referrals/Number of individuals interested in referrals                                                                                                                                                                                                            |
| Participant satisfaction                                                                                                                              | 1. Number of individuals satisfied or very satisfied with session attended or participated in/Number of individuals participating in career or education sessions<br>2. Number of individuals satisfied or very satisfied with services offered/Number of individuals interested in using services |

**Table 5.5**  
**Outdoor Recreation Equipment Checkout and Equipment Rental Capability MOPs**

| Measure of Performance                                    | Definition                                                                                                                                                                                                                                                                                |
|-----------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Inputs/Resources</b>                                   |                                                                                                                                                                                                                                                                                           |
| Facility is fully staffed                                 | Number of staff positions filled/Number of staff positions required                                                                                                                                                                                                                       |
| Staff are certified/trained                               | Number of staff with required credential or certification/Number of staff requiring credential or certification                                                                                                                                                                           |
| Facilities are maintained to standards                    | 1. Number of facility-related issues per year<br>2. Number of facility-related issues addressed in timely manner/Number of facility issues raised (in a set time frame, such as 6 months)                                                                                                 |
| Equipment has required registration/licensing             | Number of pieces of equipment that had continuous registration or licensing/Number of pieces of equipment required to have continuous registration or licensing                                                                                                                           |
| Availability of required/requested equipment              | Number of available equipment pieces/Number of required or requested equipment pieces, by equipment type                                                                                                                                                                                  |
| <b>Activities/Outputs</b>                                 |                                                                                                                                                                                                                                                                                           |
| Proportion of required hours provided                     | 1. Number of weekly hours of operation/Number of operating hours required by week<br>2. Number of weekly hours per activity type/Number of operating hours required for each activity type; calculated by activity (e.g., unit fitness or recreation programs)                            |
| Skills and safety lessons provided                        | Number of skills and safety lessons provided                                                                                                                                                                                                                                              |
| Equipment checkouts, per individual                       | Number of checkouts by individuals, by type, cost, and value of equipment                                                                                                                                                                                                                 |
| Equipment checkouts, per unit/command                     | Number of checkouts by units or commands, by type, cost, and value of equipment                                                                                                                                                                                                           |
| Utilization of skills and safety lessons                  | 1. Number of individuals receiving skills and safety lessons<br>2. Number of individuals receiving skills and safety lessons/Number of individuals required to receive skills and safety lessons                                                                                          |
| Satisfaction of Marines using facilities and equipment    | 1. Number of Marines satisfied or very satisfied with facilities and equipment utilized/Number of Marines using facilities and equipment<br>2. Number of Marines satisfied or very satisfied with facilities and equipment offered/Number of Marines using facilities and equipment       |
| Satisfaction of dependents using facilities and equipment | 1. Number of dependents satisfied or very satisfied with facilities and equipment utilized/Number of Marines using facilities and equipment<br>2. Number of dependents satisfied or very satisfied with facilities and equipment offered/Number of Marines using facilities and equipment |

**Worksheet 5.1.**  
**Operationalizing Measures of Performance**

| Input/Activity/<br>Output                                                                      | Measure of<br>Performance                             | Measure Definition                                                                                                                    | Metric or Time<br>Frame                                                  | Data Source                                                                                                  |
|------------------------------------------------------------------------------------------------|-------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------|
| (L.I.N.K.S.)<br>Sessions on<br>military culture                                                | Session<br>availability                               | Number of<br>sessions taking<br>place                                                                                                 | Calculate for one-<br>year period                                        | L.I.N.K.S.<br>registration log                                                                               |
|                                                                                                | Session<br>attendance                                 | Number of<br>families that<br>attend a session                                                                                        | Calculate for one-<br>year period                                        | L.I.N.K.S.<br>registration log,<br>in-session sign-in<br>sheets                                              |
| (FMEAP)<br>Participant<br>satisfaction                                                         | Participant<br>satisfaction with<br>sessions attended | Number of<br>individuals<br>satisfied or very<br>satisfied with<br>session/Number<br>of individuals<br>participating in<br>session    | Assess at end<br>of each session,<br>track trends over<br>6-month period | FMEAP counselors<br>to administer<br>one-question<br>satisfaction survey<br>at the conclusion<br>of sessions |
| (Outdoor<br>Recreation<br>Equipment<br>Checkout and<br>Equipment<br>Rental) Qualified<br>staff | Staff are certified/<br>trained                       | Number of staff<br>with required<br>credential or<br>certification/<br>Number of<br>staff requiring<br>credential or<br>certification | 100% of staff<br>should be<br>qualified; assess<br>every 6 months        | Personnel files                                                                                              |
|                                                                                                |                                                       |                                                                                                                                       |                                                                          |                                                                                                              |

## Defining Measures of Effectiveness

MOEs measure short-term and intermediate outcomes. They help you to determine whether your PSC is achieving its intended goals. For the MF Division PSCs, the desired outcomes generally fall into one of four broad categories: knowledge, skills, attitudes, and behaviors. For example, MOEs related to the acquisition of new *knowledge* may include a better understanding of the USMC mission, learning best practices in nutrition and exercise, or knowledge of employment options. *Skills* might include communication skills, positive coping, financial literacy, or job skills. MOEs related to *attitudes* might include marital satisfaction, self-efficacy, or social cohesion. Finally, *behaviors* might include reducing negative behaviors (e.g., substance misuse) or building a social network. As with MOPs, your logic model can be used as a tool to identify MOEs.

### Measure Considerations

There are a number of factors to consider when identifying MOEs. In this section, we review four considerations to keep in mind.

1. **Identifying and selecting a data source.** As with MOPs, you may opt to use primary data or secondary data for your MOEs. In the context of MOEs, primary data could include sources such as surveys, behavioral health measures, or knowledge tests. There may be secondary data sources relevant to your program as well. However, it is important to understand the limitations of secondary data as they relate to your MOEs. For example, the previously mentioned Marine Corps Quality of Life Survey includes broad questions about quality of life, which may be a relevant outcome for your program. However, because these measures are general and collected force-wide, it would be very challenging to attribute any improvements in quality of life to your specific PSC. As another example, programs such as the Alcohol Screening Program are designed, in part, to increase safety. One way to measure this would be to collect data about safety incidents where at least one person involved had a blood alcohol concentration above 0.00. These data may capture some of the more serious incidents but may not provide a comprehensive picture of all of the safety incidents associated with alcohol misuse. It may also be difficult to interpret nuances in the data—for example, were these on-the-job incidents, or more general incidents that took place on an installation? It may be that secondary data provide an appropriate source of data, but you will have to interpret the outcomes in the context of any limitations to the data. Triangulating different sources of data can help with such interpretations.
2. **Identifying and establishing metrics or expected patterns of performance.** As you identify MOEs for your program, consider whether there is a particular metric or pattern of performance in mind. As with MOPs, some MOEs are designed to be measured over time. You may have certain trends in mind that you would like to observe—for example, you may want to be sure there are not substantial decreases in an MOE over time—but without a specific metric to achieve. In these cases, the PSC should determine how often it wants to examine the outcome (e.g., quarterly, every six months, every year).

For other MOEs, a PSC may want to establish a particular metric. For example, for FMEAP, one MOE is the proportion of Marine family members who have a completed resume as a result of the service. FMEAP may determine that it wants at least 80 percent of family members to successfully develop a resume. As another example, one goal of L.I.N.K.S. is for family members to understand and feel confident making use of certain resources available to them, such as Family Readiness Officers. If participants are asked to rate how confident they feel accessing a certain resource on a scale of 0 to 100, L.I.N.K.S. may set a standard that all responses should reach a threshold of 75 or higher.

Finally, some measures will be examined relative to a comparison group. This is the case in experimental and quasi-experimental evaluations. If you are measuring your outcomes relative to a comparison group, you may expect that there will be a difference between the two groups. For example, you might expect that individuals who use the Outdoor Recreation Equipment Rental capability will report greater community connectedness than those who do not.

3. **Developing measures based on the content of a program.** When a PSC offers workshops, courses, or sessions on a specific topic, the primary short-term outcome is often increased knowledge specific to that topic. In these cases, the best way to measure changes in knowledge might be to develop a measure specific to the content of that session.

To develop a knowledge test, you can use materials that govern the program content. For example, many workshops may have specific learning objectives. You could develop knowledge items that relate specifically to these objectives. When programs follow a specific curriculum, the curriculum can also be used as a source to develop knowledge items. For example, the Life Skills program offers emergency preparedness workshops. This workshop covers information about natural disasters and manmade hazards and actions that families can take if these incidents occur. At the end of the session, participants could complete a knowledge test that covers the different types of emergencies and assesses for the types of emergency plans described in the workshop. As you develop a knowledge test, you can ask yourself, “What do I expect participants to have learned as a result of this course?”

4. **Using research-validated instruments to evaluate outcomes.** When selecting measures, one option is to use existing instruments that have been developed and validated through previous research or evaluation projects. For example, one outcome associated with L.I.N.K.S. is improved perceptions of social support from other military families. There are a number of instruments that have been developed to measure social support, including the Reduced Social Relationship House Questionnaire (Gottman, 1994; Gottman, 1999). Rather than developing a new measure of social support specific to the workshop, L.I.N.K.S. staff could instead administer this scale.

Often, these existing scales are designed to measure attitudes or behaviors. Examples of attitudes include perceptions of social support, feelings of depression, perceptions of quality of life, unit cohesion, or marital satisfaction. Examples of behaviors include frequency of alcohol use, interpersonal communication skills, or stress management practices. If your PSC affects these types of outcomes, you may be able to find an existing scale.

Before looking for a validated instrument, it can be useful to understand what is meant by “validity.” In this context, we are referring to instruments validated by previous research, not to instruments approved by the Marine Corps. In a research context, validity has several different facets. First, you want to make sure that you select an instrument with *construct validity*. This refers to whether the instrument measures what it intends to measure (Westen and Rosenthal, 2003). For example, does a depression instrument actually measure depression, or does it measure something else, such as general distress or anxiety? Studies look at this in different ways. For example, they might look to see whether scores on one depression measure are correlated with scores on other well-validated depression measures (*convergent validity*). They might also examine the association between scores on the depression measure and scores on other types of measures, like anxiety measures, with the goal of making sure that the association is not too strong (*discriminant validity*). You might also look for instruments that have demonstrated *criterion validity*—that is, scores on the instrument are related to a relevant outcome or other construct of interest (Westen and Rosenthal, 2003). It is also important to know that the validity of a measure depends on the specific sample on which it was validated. For example, it may demonstrate good validity in female civilians in the

general community but have different validity in male service members. When you find an instrument with established validity, it can be referred to as *evidence-based*—that is, it has its foundation in scientific research.

You may be wondering how to identify these types of existing scales. A good place to start in finding these scales is an internet search. For example, if your program targets interpersonal communication skills, you might search for “measure” or “instrument” plus “interpersonal communication skills.” Alternatively, if you know of any research studies that have examined a PSC like yours, you may look to that study to see what type of instrument was used. If you are familiar or comfortable with academic databases, you might try searching databases such as Google Scholar. If you are not familiar with these databases, you can enlist the librarian at an installation library to assist with this search.

In many cases, there might be multiple instruments designed to measure your outcome of interest. As you review the available options, there are a few questions to keep in mind.

- Do the studies describe the validity of the measure? Ideally, you will be able to find a measure with good construct validity, and potentially with good criterion validity as well.
- Has the instrument been used in military populations before? If so, you may have more confidence that it is an appropriate choice for your PSC. If the instrument has been used only in nonmilitary populations, it may perform differently in your population.
- How many studies have used the instrument? Some instruments have been constructed and examined for a single study, whereas others have been used or validated by multiple studies. If given the choice, it is generally better to identify an instrument used and validated in more studies.
- Is the length of the instrument appropriate for your purposes? Some instruments that are used in research or clinical settings can be quite long, which may be overly burdensome for your program participants. You may find shorter measures by searching for “short forms” or “brief measures.”
- Does the content of the items match your PSC? You may be looking for a measure of social support but then discover that the specific items on some of the scales that you are finding do not quite relate to the activities you offer. Take a read through any measure first to make sure that the questions fit with your PSC; otherwise, you may not detect the effects of your PSC.

Sometimes, you might find an instrument that seems like a fairly good fit, but the wording of the questions is not quite right. For example, maybe the measure would be more appropriate if it used military-specific language, or if certain items were removed or added. In this case, you can use the items on the measure as a starting place but then make modifications as necessary: adjust the language, add items, delete items. One word of caution is that when you modify measures for your own purpose, it can affect their validity. For example, when you make changes to the wording or content, it is possible that the instrument will have different construct validity—that is, maybe it will not assess the target outcome in the same way. This type of modified measure may be

**Table 5.6**  
**Sample L.I.N.K.S. Program MOEs**

| Measure of Effectiveness                                                                 | Definition/Data Source                                                                                                                    | Notes for Use                                                                                                                                                                 |
|------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Short-Term Outcomes</b>                                                               |                                                                                                                                           |                                                                                                                                                                               |
| Increased basic military knowledge (e.g., deployment, rank structure, chain of command)  | L.I.N.K.S. Challenge Quiz                                                                                                                 | Compare quiz performance before and after programming.                                                                                                                        |
| Increased understanding of USMC mission, organization, and culture                       | L.I.N.K.S. Challenge Quiz                                                                                                                 | Compare quiz performance before and after programming.                                                                                                                        |
| Ability to understand a Leave and Earnings Statement                                     | Quiz should be developed regarding specific program learning goals related to understanding military pay.                                 | Compare quiz performance before and after programming.                                                                                                                        |
| <b>Intermediate Outcomes</b>                                                             |                                                                                                                                           |                                                                                                                                                                               |
| Understanding of family readiness, deployment, and separation                            | Quiz should be developed regarding specific program learning goals related to understanding family readiness, deployment, and separation. | Compare quiz performance before and after programming.                                                                                                                        |
| Increased social support from other military families                                    | Multidimensional Scale of Perceived Social Support (Zimet et al., 1988)                                                                   | Items should be adapted to refer specifically to military families. Compare responses before and after programming.                                                           |
| Improved ability to activate available resources (e.g., Family Readiness Officers, MCCS) | Domain-specific self-efficacy (Bandura, 2006), measured for each relevant resource                                                        | Item should be adapted to reflect participant confidence to activate specific resources. Compare responses before and after programming.                                      |
| Improved relationship skills                                                             | Reduced Sound Relationship House Questionnaire (Gottman, 1994; Gottman, 1999)                                                             | Compare responses before and after programming. Post-program measure should be administered after a delay to be sensitive to change (e.g., 6 months following participation). |

NOTE: This program serves a range of USMC family members, including spouses, parents and extended family members, teens, and children. The content and format of the L.I.N.K.S. programming may vary depending on the specific target population. The MOEs in this table are for military spouses. The MOEs and/or measure definitions may change based on the population (e.g., use of measures developed and/or validated for youth).

referred to as “evidence-informed”—that is, it is based on a scientifically valid foundation but has been altered or tailored to your PSC. It may still be an appropriate measure for your program, but this is a limitation to keep in mind.

### Sample MOEs

In Tables 5.6, 5.7, and 5.8, we identify MOEs for the three sample PSCs. In the first column is the MOE; in the second column is the way each MOE is operationalized and/or the data source (e.g., if a specific existing instrument is being used); and in the third column are additional notes for use. The information in this third column often addresses the timing of these measures; for example, these notes describe when a measure should be collected before and

**Table 5.7**  
**Sample FMEAP Service MOEs**

| Measure of Effectiveness                                                                                                        | Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Notes for Use                                          |
|---------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------|
| <b>Short-Term Outcomes</b>                                                                                                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                        |
| Family members follow up on referrals/resources provided by program staff                                                       | Number of referred individuals who engaged or used resources/Number of individuals who received a referral                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Measure over time                                      |
| Increased knowledge or skills regarding career-related topics (e.g., resume content, conducting a job search, interview skills) | Quiz should be developed regarding specific topics covered in career training/education workshops                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | Compare quiz performance before and after programming. |
| Completed resume                                                                                                                | Number of family members receiving career coaching who have a completed resume/Number of family members receiving career coaching who need to develop a resume                                                                                                                                                                                                                                                                                                                                                                                                                                                              | Measure over time                                      |
| Prepared for job interviews                                                                                                     | Number of family members who report feeling prepared for a job interview/Number of family members receiving preparation for job interviews                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Measure over time                                      |
| <b>Intermediate Outcomes</b>                                                                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                        |
| Measurable progress toward career goals                                                                                         | <p>Specific measures will depend on presenting issues raised during career coaching. Sample measures may include</p> <ol style="list-style-type: none"> <li>1. Number of job applications resulting in interviews for family members using FMEAP/<br/>Number of applications submitted by family members</li> <li>2. Number of job applications resulting in job offers for family members using FMEAP/<br/>Number of applications submitted</li> <li>3. Number of family members who use FMEAP to apply to educational programs/<br/>Number of family members who use FMEAP to pursue educational opportunities</li> </ol> | Measure over time                                      |

after a session, or when there should be a delay between the end of the activity and the measurement of the outcome (e.g., six months later).

A few notes about these tables. First, as with the MOPs, the MOEs are directly tied to the logic models presented in Chapter Three. These MOEs reflect ways to measure the short-term and intermediate outcomes of each activity.

Second, you will notice that the types of MOEs differ quite a bit, depending on whether the MF offering is a program, service, or capability. For example, L.I.N.K.S. is a well-defined program with a specific curriculum. This makes the selection of MOEs easier, because the program curriculum is designed to impart certain knowledge and skills to participants. Some of the outcomes of the program are very specific to L.I.N.K.S. or to the USMC—for example, an under-

**Table 5.8**  
**Outdoor Recreation Equipment Checkout and Equipment Rental Capability MOEs**

| Measure of Effectiveness                         | Definition                                                                                                                                                                                                                                                                                                        | Notes for Use                                                                                                                                                                                                                                                                                                                          |
|--------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Short-Term Outcomes                              |                                                                                                                                                                                                                                                                                                                   |                                                                                                                                                                                                                                                                                                                                        |
| Access to affordable/healthy leisure activities  | Average expenditures on outdoor equipment for those who use Outdoor Recreation Equipment Rental                                                                                                                                                                                                                   | Measure over time                                                                                                                                                                                                                                                                                                                      |
| Skills for using outdoor equipment               | Improved knowledge, skills, or abilities related to outdoor equipment. Quiz or measure should be developed based on specific content of skills training.                                                                                                                                                          | Measure over time                                                                                                                                                                                                                                                                                                                      |
| Intermediate Outcomes                            |                                                                                                                                                                                                                                                                                                                   |                                                                                                                                                                                                                                                                                                                                        |
| Increased quality of Marine Corps life           | Ratings of quality of life among those who use Outdoor Recreation Equipment Rental. Item should be specific to use of this capability. For example, "How much does having access to these resources contribute to your quality of life as a military family?" Scale could range from "Not at all" to "Very much." | Measure over time                                                                                                                                                                                                                                                                                                                      |
| Increased perceptions of community connectedness | Ratings of community belonging. Sample measures for use include Military Family Institute Base Cohesion (McClure and Broughton, 2000) or Brief Sense of Community Scale (Peterson, Speer, and McMillan, 2008).                                                                                                    | Items should be adapted as necessary for the USMC community, particularly the Brief Sense of Community Scale, which is not specific to military. Measure as a comparison between individuals who use Outdoor Recreation Equipment Rental and a similar group of individuals (Marines and/or dependents) who do not use the capability. |
| Improved unit cohesion                           | Ratings of unit social cohesion. Sample measures include Group Cohesion–Social Scale (Chang and Bordia, 2001) or Cohesion Questionnaire–Company Survey Social Scale (Smith and Hagman, 2004).                                                                                                                     | Measure over time. Scales may need to be tailored to be specific to the USMC community, particularly the Group Cohesion Scale, which is not specific to military.                                                                                                                                                                      |

standing of USMC career paths or USMC volunteer and community involvement opportunities. For these outcomes, creating a knowledge quiz specific to the content of the L.I.N.K.S. workshop is probably the best way to measure whether participants gained this knowledge. In addition, some of the outcomes relate to broader psychosocial constructs, such as perceived social support or relationship skills. For these types of outcomes, there are many existing scales that can be used or adapted. We selected two existing instruments for these MOEs. Our choices were based on the specific items on each scale and on the length of each scale: We wanted to select instruments that were relevant to the outcome but not overly burdensome for program participants. In the "Notes for Use" column for these two measures, you can see that we recommended ways that these existing instruments could be adapted to fit even better with the L.I.N.K.S. curriculum.

Other support services and capabilities provide information to Marines and family members, often passively. The goal might be to have information available to meet the needs of the individuals who access the service or capability; however, the specific needs will vary by indi-

**Worksheet 5.2****Operationalizing Measures of Effectiveness**

| Measure of Effectiveness                                                                           | Measure Definition                                                                                                                                                                               | Notes for Use<br>(e.g., comparison group, time frame for data collection/analysis, metrics)                                                                                                                                                                             |
|----------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| (L.I.N.K.S.) Basic knowledge about USMC and military lifestyle                                     | Scores on L.I.N.K.S. Challenge Quiz                                                                                                                                                              | Measure at beginning and end of L.I.N.K.S. workshop. Examine change from pre- to post-workshop.                                                                                                                                                                         |
| (FMEAP) Marine family members have a completed resume                                              | Proportion with completed resume: Number of family members receiving career coaching who have a completed resume/Number of family members receiving career coaching who need to develop a resume | Should be measured over time. Examine trends in this proportion at 6-month intervals. Target is to maintain at least 80-percent proportion over time.                                                                                                                   |
| (Outdoor Recreation Equipment Checkout and Equipment Rental) Improved quality of Marine Corps life | Single quality of life item: "How much does having access to these resources contribute to your quality of life as a military family?" 5-point scale ranging from "Not at all" to "Very much."   | Should be measured over time among those individuals who use the service. Could be broken down separately for Marines and dependents, or for individual Marines who use the service versus those who use it as part of their unit. Examine trends at 6-month intervals. |
|                                                                                                    |                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                         |

vidual. In this case, the MOE will likely pertain to whether the service or capability met the needs of its users. For example, family members who use FMEAP may present with a variety of career-related goals. They might be interested in developing a resume, or need help with interviewing skills, or want to explore advanced educational opportunities. Within these broad categories, though, their specific needs might vary. Therefore, you would not want to provide all participants with a knowledge test measuring resume skills, as this may not adequately reflect the services they received. A more appropriate way to define this measure might be to calculate the proportion of family members who needed to develop a resume who did so during their encounters with FMEAP.

Because this MOE is defined using a proportion, it may look more similar in format to the MOPs described above. However, the key difference is that the MOPs were designed to do things like track the number of individuals *participating* in services, or the proportion of interested parties who *received* services. Those measures are focused on usage. By contrast, the

proposed MOEs measure the outcome of the career training or education—in this case, the successful completion of a resume.

Finally, the MOEs for Outdoor Recreation Equipment Rental provide an example of measures that would ideally be examined in a comparison group. Many aspects of the MF Division programming—and of the broader set of resources available to Marines and their families—are designed to promote community connectedness and/or unit cohesion. Therefore, to best determine whether this specific capability contributes to those outcomes, it would be helpful to measure the same outcomes in a group of Marines and/or dependents who have not used the capability. If you found that community connectedness or unit cohesion was substantially higher among individuals who used this capability than among those who did not, you would have more confidence concluding that those outcomes were due to the activity and not some other program or service.

### **Identifying MOEs for Your Program**

Using this information from your logic model and the information reviewed in this chapter, you can use Worksheet 5.2 to define your MOEs. We have provided examples based on our three sample programs and the MOE lists earlier. In the first column, identify your MOE. Remember that MOEs correspond to your program's short-term and intermediate outcomes, which you can find on the logic model you created. In the second column, indicate the specific definition for that measure. Will you use an existing instrument or scale? Create one for your activity? Calculate the proportion of participants who achieve a specific outcome? Finally, in the third column, identify some specific notes for use. For example, you might indicate when the measure should be administered or calculated, or how the measure will be used and interpreted (e.g., whether it will be examined over time, compared with a specific metric).

### **Summary**

This chapter provided information about the difference between MOPs and MOEs, how to identify these measures, and considerations to take into account when figuring out how to define these measures. After using this chapter, you should have identified MOPs and MOEs for your evaluation plan (Worksheets 5.1 and 5.2). Now you are ready to start implementing your evaluation plan. The next chapter describes the next stages of your evaluation: data collection, analysis, and interpretation.



## Conducting Your Evaluation

---

This chapter begins by describing how to translate plans for evaluating a program, service, or capability (PSC) into the actual implementation of an evaluation. We describe potential tools or processes to collect and organize the data needed for an evaluation and establish a database. Next, we review specific types of data and analysis and orient you to a tool in Microsoft Excel that the RAND research team developed to support this analysis. Finally, we describe approaches to interpreting analyses to draw conclusions about trends over time, performance of the PSC, and effectiveness of the PSC. We discuss key considerations to take into account in order to draw accurate conclusions in this process, based on the strengths and limitations of different evaluation designs.

### Implementing an Evaluation Plan

There are several key decisions in translating an evaluation plan to implementation of an evaluation, including determining who will conduct the evaluation and over what time frame, and with what existing or new data sources.

#### Who Will Conduct the Evaluation?

An evaluation may be conducted by PSC managers and staff, who can incorporate new data collection and/or make use of data collection that already takes place. Other staff may also be a critical part of data collection efforts for MOPs and MOEs, by doing such tasks as documenting attendance at workshops, directing USMC family members to a feedback survey to complete, recording referrals that they make, or providing information about their credentials.

The Excel tool developed by RAND is intended to make it straightforward and accessible for managers and program staff to analyze MOP and MOE data and draw conclusions. More advanced analyses beyond the scope of this chapter can be conducted using an Excel add-on or a statistical software package (e.g., R). In some cases, consulting an expert in evaluation or statistics may be valuable for implementing large-scale data collection, conducting more advanced analyses, or addressing irregularities in a data set.

#### Over What Time Period Will the Evaluation Take Place?

As referenced in previous chapters, the time course of an evaluation can also vary. Data collection and analysis can take place at specific time intervals (e.g., quarterly, yearly) or in relation to specific sessions or events (e.g., a workshop). Time intervals may be selected based on when

and across what period of time specific changes are expected to take place, such as changes in program content or changes in the lives of participants.

For example, for the FMEAP service, an MOE of family members feeling prepared for job interviews could be assessed with each individual directly after participating in an activity such as career coaching (e.g., through a feedback survey that the counselor invites the participant to complete in a private room and submit before leaving). However, the FMEAP MOE of progress made toward career goals would need to be assessed at a later point, to allow individuals time for any changes in progress toward career goals that may take place. An option in such a case could be emailing a link to an online survey, on a rolling basis over a period such as six months, to individuals who participated in career coaching within that period (perhaps replacing the immediate feedback survey).

## Designing Processes to Collect Required Data

The RAND tool is designed to help you organize, track, and understand your data. For most MOP and MOE data collected as part of your evaluation, you will be able to copy your data from the original source and paste it into the tool, which will output descriptive information about the data and the results of relevant statistical analyses. The tool is primarily oriented around individual-level data. This means that information would be collected and stored for each person or family of interest (i.e., each person's or family's data are stored in a separate row in an Excel spreadsheet). But elements of the tool will also process data collected and stored at the aggregate level (e.g., the number of people who attended a workshop, the number of trained staff). This section details the collection and storage of individual versus aggregate data and how to transfer your data into the RAND tool.

### Ways of Collecting Data from Participants at the Individual Level

In the previous chapter, we identified individual-level surveys, tests, and other instruments as a common source of data. When you created your evaluation plan, you considered the timing of administration and sample for each instrument, as well as some of the staffing and infrastructure needs associated with data collection.

It is also useful to think through the procedures for actually collecting data at the individual participant level. The following are suggestions for ways to collect individual-level data via surveys or other individual-level instruments (e.g., knowledge tests, behavioral measures):

1. **Hand out short paper surveys at the point of service use.** For example, you might distribute surveys after a workshop, training, or meeting to people actively using a service (e.g., at the library, at the gym). To ensure maximum participation, make sure that the surveys are short—no more than one to two pages double-sided. Following are some tips on keeping secure the personal information of those who complete paper-based surveys.
2. **Gather email addresses or other contact information for PSC users and follow up with an online survey invitation afterward.** If your PSC already collects contact information from users, it is good practice to notify users that you will be contacting them for the survey (e.g., during a visit, tell the user, “We are conducting a survey to get feedback on how our program is working for you. You will be receiving an email

from us in the coming week asking you to complete a short survey about your visit”). If your PSC does not collect contact information from users, provide a way for them to give you that information using a private and secure method. For example, you could have users write their email addresses on cards and drop them into a secure box, to be entered into a contact database for the survey. However the contact information is collected, it is important to protect the confidentiality of users by securely storing this information in a manner that limits access to only those who need it to invite them to take the survey. Regarding the actual data collection instrument, you can set up an online survey free. With limited exceptions, the use of commercially available internet-based capabilities is prohibited (Department of Defense Instruction 8550.01, 2012). Be careful when considering which online service to use. Popular services, such as SurveyMonkey and REDCap, are internet-based capabilities that are subject to DoDI 8550.01 and are *not* approved. The Office of Management and Budget (OMB) developed an alternative online capability, MAX Survey, for use by federal agencies. MAX Survey is a powerful, user-friendly tool containing robust options for navigation and within-survey conditional logic. To access and set up MAX Survey, visit the MAX.gov homepage: <https://max.omb.gov/maxportal/home.do>.

Whether a survey is implemented with participants on paper or online, it is important to make respondents feel comfortable giving honest answers and feedback. Letting them know that their responses will be confidential is one way to do so. To allow respondents to feel comfortable providing suggestions or expressing dissatisfaction, it can also be valuable to let them know that feedback is being solicited to understand users’ experiences and improve the resources made available.

In the following box, we elaborate on suggestions for implementing surveys to help achieve your intended sample size. The process for collecting data should ideally remain the

#### Considerations to Increase Survey Response Rates

- **Consider timing.** Try to avoid implementing a survey during times of the year, week, or day when individuals might be less likely to respond. For example, holidays may be times when people are busy and unlikely to respond. Individuals may also be more likely to miss emails that come late in the day.
- **Send reminders.** Reminders can help to get additional respondents, given that individuals are busy and may forget to respond. However, there are diminishing returns to additional reminders, so it may not be useful to send out more than three or four. It is also important to not annoy individuals with reminders, as this may reduce future response rates. If possible, avoid sending reminders to those who have already completed the survey.
- **Give individuals a reason to complete the survey.** Telling individuals why the survey is being fielded and how their responses will be used to inform PSC practices can provide a greater incentive to complete the survey. It may also be useful to appeal to other values, such as supporting family time or contributing to overall military family quality of life.
- **Limit the length and number of surveys.** The more time a survey takes to complete, the less likely an individual may be to respond. Keeping surveys short and informing individuals about the minimal requirements may be useful in persuading individuals to respond. In addition, overburdening individuals with multiple surveys may diminish response rates.
- **Provide incentives.** Providing small monetary or nonmonetary rewards (e.g., pens, gift certificates, candy) for completion of a survey can help to encourage survey completion.
- **Consider anonymity.** It may be useful to allow for anonymous responses so that individuals are not prevented from responding by concerns about risk or privacy. On the other hand, other populations may see a lack of anonymity as a benefit to express their feedback, so it is important to understand the population and the topics being covered within the survey.

SOURCE: Daugherty et al., 2016.

same across different time points of data collection to make the responses from these different time points as comparable as possible.

### **Considering Data Collection from Nonparticipants**

In some cases, you may want to collect data from nonparticipants to compare their responses on MOEs to those of participants (e.g., if you are using a quasi-experimental design). You will want to collect data from individuals who are as similar as possible in all ways to participants so that they can serve as a good comparison. Preexisting differences between nonparticipants and participants, such as living farther away or being less involved on the installation, make it harder to be sure that any differences between them are based on use of the PSC. Note that it is impossible to rule out all such differences, and thus comparison between nonparticipants and participants cannot be said to show causal *effects* of the PSC.

One option, where possible, is to conduct a survey with individuals who have expressed interest in the PSC but have not yet been able to use it. For example, you could reach out to individuals on a wait list for a workshop, or who are signed up for an upcoming workshop, and ask them to complete a survey.

Responses from nonparticipants can be included in a database at the individual level in the same way as responses from participants. You will want to include a group membership variable (discussed more below) indicating who was a participant and who was not. We continue in this chapter referring to data from participants, but with the understanding that in some cases data may be from nonparticipants.

### **Documenting Data Collection Processes**

Whether existing data are analyzed retrospectively or new data are collected, it is important to keep clear records of how the data were gathered so that the procedure can be replicated in the future. For example, for the L.I.N.K.S. program, one MOP is the frequency of updates to session content. To track this MOP, relevant staff members may need to be trained such that if they make an update to the session content, they also mark down this update (e.g., date, staff member name, summary of updates made) in a spreadsheet or another shared document. It would be important to document what type of update occurred each time that someone marked down that an update had taken place, to be consistent in the definition of an “update.” Consider the possibility that at one time, someone recorded an update only if a new activity was added to a workshop, but at another time, someone recorded an update if any text or images were modified in a slide. The different definitions used would make the different data points less comparable and would make it difficult to interpret any trends in the frequency of updates over time.

Organizing and labeling data, and documenting how data points were defined and collected, is helpful both for when the data are analyzed and for when collecting the same data point again in the future. One easy way of documenting these definitions and collection processes is to leave comments on cells that contain column titles in an Excel spreadsheet. Another option is to create a separate spreadsheet tab within a database to record information about when, where, and how the data were collected.

## Designing Tools to Collect Required Data<sup>1</sup>

### Maintaining Databases at the Individual Versus Aggregate Level

The structure of your database will depend on the nature of the evaluation for your PSC, and it is likely that multiple databases may be needed for different data collection efforts or points in time. For example, for the L.I.N.K.S. program, there is likely to be data collection with individual session participants for different types of sessions, as well as data collection at the aggregate level about the availability, attendance, and staffing of sessions. Data sets for different cohorts of participants, for specific workshops or sessions rather than a PSC as a whole, or for a PSC as a whole rather than individuals, are best kept as separate databases.

### Developing a Database at the Individual Level

An evaluation database is structured to contain rows and columns of data. For information collected from (or about) individual participants or families, each row represents a different PSC participant, and each column represents a different response from (or information point about) each participant or family. The columns are called *variables*.

### Unique Identifiers

To protect the privacy of participants, a unique identifier should be used for each participant in the database, in place of participant names and other identifying information (e.g., street address). The first time that a PSC interacts with or collects data from a set of participants, each participant should be assigned a unique identifier in the form of a numeric code (e.g., 101, 102, 103 . . .). This assignment could take the form of prewriting a numeric code in the corner of each paper-based survey in a stack before the stack is passed out to participants, or prewriting a numeric code as a column of an attendance sheet next to each participant's name.

Once data collection has taken place and the unique identifiers are incorporated in the database, names can be removed from the database and kept in a single, secure location (e.g., a locked cabinet or password protected electronic file) along with the information of which unique identifier corresponds to each name. To comply with IRB-approved requirements, this file must be stored in a separate location from any other data from participants (e.g., a separate file cabinet, a different computer or server). The data analysis then moves forward using only the numeric values rather than the names and identifiable information. The database linking unique identifiers to participant names must be stored for the entire data collection process and not discarded until after data collection and analysis are completed. This allows for the same identifier to be used for each participant when a second round of data collection takes place with the same set of participants.

### Entering Data by Hand

When data are entered in Excel by hand, such as from a paper-based survey, each value corresponding to a participant's response is entered in the corresponding column of that participant's row in the database. Several steps can reduce human error during manual data entry. First, columns can be specified in Excel as requiring a specific information format (e.g., numeric value, date, text) or even a specific range of values. The conditional formatting feature allows for

<sup>1</sup> Two policies are relevant to this discussion: (1) Secretary of the Navy Manual 5210.1 Chapter 1, *DON Records Management Program Records Management Manual*, September 15, 2017; and (2) Secretary of the Navy Instruction 5211.5E, *DON Privacy Act Program*, December 28, 2005. Please review them prior to any actual data collection efforts begin.

highlighting values that do not fall within a required range (e.g., between 1 and 5), as shown in Figures 6.1 and 6.2. This automation can improve the accuracy of the data entry.

Second, having multiple staff members involved in data entry allows for catching and remedying errors. A second person can check all values entered by a first person, and a supervisor can check a subset of the data entries (e.g., 10–25 percent). As errors are detected, they can be addressed as a group to develop steps to prevent similar inconsistencies.

### Using Survey Scales

In some cases, as referenced in earlier chapters, an appropriate MOE may be a published survey *scale*. Some scales require purchasing an academic journal article to access the text of the survey questions that compose the scale, while others are available in public articles. Some scales also require a fee for each respondent to whom they are administered.

A scale is composed of a set of multiple, related questions for participants, rather than a single question. The advantage of having multiple related questions is that the MOE does not rely on the particulars of any single question, and a more reliable, averaged score can be obtained for each person across responses to different questions.

**Figure 6.1**  
Sample Input to Use Conditional Formatting to Identify Values out of Range

RAND TL264-6.1

**Figure 6.2**  
Sample Output of Conditional Formatting to Identify Values out of Range

| Participant | Scale item 1 | Scale item 2 | Scale item 3 | Scale item 4 |
|-------------|--------------|--------------|--------------|--------------|
| 1001        | 4            | 5            | 4            | 5            |
| 1002        | 4            | 4            | 3            | 9            |
| 1003        | 5            | 5            | 5            | 5            |
| 1004        | 4            | 5            | 0            | 5            |
| 1005        | 4            | 4            | 4            | 4            |
| 1006        | t            | 4            | 3            | 4            |

RAND TL264-6.2

### Preparing to Use a Scale

In some cases, a subset of items on the scale may need to be selected, particularly if the scale is too long to administer. The subset can be selected based on the questions' relevance to the PSC's intended outcomes and/or representativeness of the scale's different types of items. In some cases, scales have been analyzed statistically in an article to understand their different *factors*. A relevant factor within a larger scale could be selected without using other factors in the scale.

The questions in a survey scale may also need to be adapted for use. Minor modifications to scales, such as changing a single word or phrase for relevance, can be implemented and should ideally be followed up with a reassessment of the scale's psychometric properties. A statistician or a quantitative researcher with expertise in survey development or psychology should perform this assessment. If entire items are deleted from the scale or more substantive changes are made to the content, it is all the more vital to retest the scale in this way before using it as an MOE.

For example, the Multidimensional Scale of Perceived Social Support (Zimet et al., 1988) is a suggested MOE for L.I.N.K.S. that would require adapting. This scale has a set of 12 items about perceived support, with some relating to support from a significant other, some relating to support from family, and some relating to support from friends. If the scale is used as a starting point for assessing perceived support from other military families, then it could be adapted, for example, in the following way. The four items referring to "my friends" could be used and reworded to refer to "military families at my installation" or perhaps "other military families." Here, "My friends really try to help me" could become "Other military families really try to help me."

Many scales use items phrased not as questions, but rather as statements. The source of the scale should have information about a suggested way of presenting the items or type of response scale. However, for items that are statements, the response scale typically assesses the respondent's level of agreement with each statement. A scale from 1 (strongly disagree) to 5 (strongly agree) can be used for consistency across different types of measures, in the absence of reason to use a different response scale.

### Using Data from a Scale

When a survey scale is used, the response for each scale question should be entered as a separate column. The following two steps may then be needed to prepare the scale data for analysis: reverse-scoring as needed, and generating an average score. We use as an example the Anger Management Scale (Stith and Hamby, 2002), which is a suggested MOE for the Family Readiness Life Skills program, as it allows for illustrating the process of reverse-scoring.

### Reverse-Scoring

Some survey scales include questions marked as requiring *reverse-scoring*. These questions are posed in the opposite way as other questions are posed. In the Anger Management Scale, most items take a form such as, "When arguing with my partner, I often raise my voice," for which *higher* agreement on the response scale would indicate greater anger challenges. However, a subset of questions take a form such as, "I am able to remain calm and not get angry with my partner," for which *lower* agreement on the response scale would indicate greater anger challenges. Not all survey scales have questions that need to be interpreted in this reversed way.

Those that do will mark questions that require reverse-scoring, for example with the letter “R” or an asterisk.

For questions requiring reverse-scoring, the bottom end of the response scale (e.g., 1 = strongly disagree) will be treated as the top end of the scale (e.g., 5 = strongly agree), and each point on the response scale will be reversed accordingly (e.g., for a 1–5 scale: 1 converts to 5, 2 converts to 4, and 3 remains as 3). To track what changes have been made to the data set, a new column should be generated for such questions, and the original column should be left in its original form. The new column should have the same title but with the term “reverse-scored” at the end. The original column can have the term “original” at the end, for clarity.

For basic response scales with values that are all positive or zero, there is a simple way to reverse-score in Excel. In the new column, enter the numeric value that is 1 greater than the highest value on the response scale, and subtract from this value the column with the original values. For example, for a 1–5 response scale, the new column would be 6 minus the original column. An original value of 1 would then become 5 (6 minus 1), an original value of 2 would become 4 (6 minus 2), and so on. We illustrate this process in Figure 6.3. The formula written below can be entered in the top cell of the new column and then dragged down to fill the full column.

**Formula for reverse-scoring a scale item in a new column:**

= (highest value on response scale + 1) – (original response value)

### Generating an Average Score

One common way to use data from a survey scale is to create an average score for each individual, reflecting how the person tended to answer the different questions on the scale. To generate an average score, a new column should be added, titled with the name of the scale followed by the term “average score.” Each cell in this column should be calculated for its corresponding row (i.e., based on the response values for a single participant or family). The column should contain the average of the columns for the questions that make up the scale, reverse-scored if needed (i.e., only the final columns, not the original columns for questions that were reverse-scored). The AVERAGE() function can be used, with commas separating cells to be included

**Figure 6.3**  
Sample Reverse-Scoring of a Scale Question

|             | <i>When arguing with my partner, I often raise my voice.</i> | <i>I am able to remain calm and not get angry with my partner. (R)</i> | <i>I am able to remain calm and not get angry with my partner. (R)</i> |
|-------------|--------------------------------------------------------------|------------------------------------------------------------------------|------------------------------------------------------------------------|
| Participant | Scale item 1                                                 | Scale item 2 (original)                                                | Scale item 2 (reverse-scored)                                          |
| 1001        | 4                                                            | 1                                                                      | =6-C12                                                                 |
| 1002        | 4                                                            | 2                                                                      | 4                                                                      |
| 1003        | 5                                                            | 2                                                                      | 4                                                                      |
| 1004        | 4                                                            | 3                                                                      | 3                                                                      |
| 1005        | 4                                                            | 1                                                                      | 5                                                                      |
| 1006        | 3                                                            | 1                                                                      | 5                                                                      |

in the average. The formula can be entered in the top cell of the new column and then dragged down to fill the full column.

We illustrate this process in Figure 6.4, using only two scale items for simplicity to generate an average score. Note that for the item that was reverse-scored, the values from the original scale item column are not included in that respondent's average score. It is this average score variable, with a single value for each participant who answered the scale questions, that should be copied and pasted into the appropriate Excel tool to be used in analyses.

**Figure 6.4**  
Sample Generation of an Average Score for a Scale

|             | <i>When arguing with my partner, I often raise my voice.</i> | <i>I am able to remain calm and not get angry with my partner. (R)</i> | <i>I am able to remain calm and not get angry with my partner. (R)</i> |                     |  |
|-------------|--------------------------------------------------------------|------------------------------------------------------------------------|------------------------------------------------------------------------|---------------------|--|
| Participant | Scale item 1                                                 | Scale item 2 (original)                                                | Scale item 2 (reverse-scored)                                          | Scale average score |  |
| 1001        | 4                                                            | 1                                                                      | 5                                                                      | =AVERAGE(B12,D12)   |  |
| 1002        | 4                                                            | 2                                                                      | 4                                                                      | 4                   |  |
| 1003        | 5                                                            | 2                                                                      | 4                                                                      | 4.5                 |  |
| 1004        | 4                                                            | 3                                                                      | 3                                                                      | 3.5                 |  |
| 1005        | 4                                                            | 1                                                                      | 5                                                                      | 4.5                 |  |
| 1006        | 3                                                            | 1                                                                      | 5                                                                      | 4                   |  |

RAND TL264-6.4

### Using a Group Membership Variable

When data are collected for individuals from different groups or subgroups that will be compared (see next section on analysis types), a column should be added to the data set to indicate each individual's group membership. For example, if individuals who participate in Outdoor Recreation Equipment Rental will be compared with individuals who do not use this capability, the group membership column could be titled "Group: User" and the values entered in each cell could be 1 for nonusers and 2 for users (see Figure 6.5).

The data set can then be sorted by the group membership column (see Figure 6.6) such that, for example, all rows with a "1" in the group membership column are together, and all rows with a "2" in that column are together. The values for columns of interest (ID, outcome being analyzed, and group membership) for each group, separately, should then be copied and pasted on the appropriate spreadsheet tab in the Excel.

### Documenting Column Values

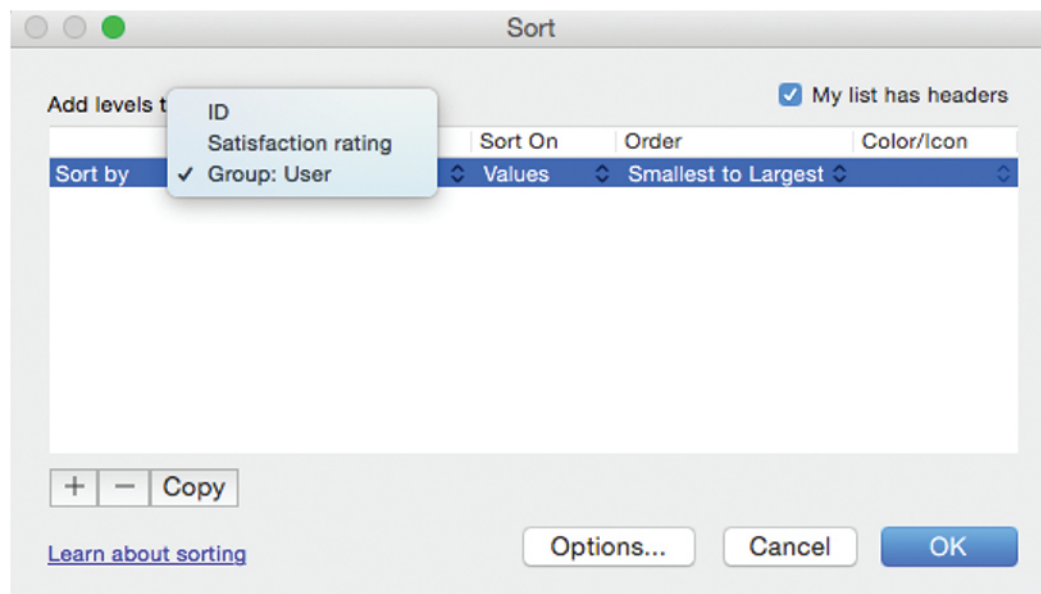
As described above, keeping records during data collection and analysis is key. One important piece of information to document is the different potential values for each variable and the meaning of these values (e.g., the labels for each scale point from 1 to 5 on a response scale). As referenced previously, one way to store this information is in Excel comments on cells that contain column titles (see Figure 6.7).

**Figure 6.5**  
**Sample Use of a Group Membership**  
**Variable in Database**

| ID  | Satisfaction rating | Group: User |
|-----|---------------------|-------------|
| 201 | 2                   | 1           |
| 202 | 3                   | 1           |
| 203 | 4                   | 1           |
| 204 | 5                   | 1           |
| 205 | 4                   | 1           |
| 212 | 1                   | 1           |
| 213 | 2                   | 1           |
| 214 | 3                   | 1           |
| 215 | 1                   | 1           |
| 216 | 1                   | 1           |
| 301 | 3                   | 2           |
| 302 | 4                   | 2           |
| 303 | 4                   | 2           |
| 304 | 5                   | 2           |

RAND TL264-6.5

**Figure 6.6**  
**Sample Use of Sorting by Group Membership Column**



RAND TL264-6.6

### Transferring Data into the RAND Tool

After creating a database of responses at the individual level, you will use columns only one at a time to conduct the needed analysis of an MOP or MOE. These analyses using the RAND tool are described in the next section.

**Figure 6.7**  
**Sample Use of Comments in Excel to Document Column Values**

The figure shows an Excel spreadsheet with three columns: ID, Satisfaction rating, and Group: User. The data rows are as follows:

| ID  | Satisfaction rating | Group: User |
|-----|---------------------|-------------|
| 201 | 2                   | 1           |
| 202 | 3                   | 1           |
|     |                     | 1           |
|     |                     | 1           |
|     |                     | 1           |
|     |                     | 1           |
|     |                     | 1           |
| 214 | 3                   | 1           |
| 215 | 1                   | 1           |
| 216 | 1                   | 1           |

Two callout boxes provide scale information:

- Satisfaction rating scale:**
  - 1 = Very Dissatisfied
  - 2 = Dissatisfied
  - 3 = Neither Satisfied Nor Dissatisfied
  - 4 = Satisfied
  - 5 = Very Satisfied
- Group: User scale:**
  - 1 = Non-user
  - 2 = User

RAND TL264-6.7

When analyzing a specific MOP or MOE, you will copy and paste the content of the needed columns, including the ID column (and a group membership column if needed), into the corresponding columns of the RAND Tool to conduct analyses. The Excel tool consists of a set of spreadsheets with tabs called workbooks. We describe throughout how to paste data into different spreadsheet tabs within a workbook to use the tool. For example, to compare between two groups, one group's data will be pasted into one spreadsheet tab, and the other group's data will be pasted into a different spreadsheet tab, within a single workbook. Data collected at different points in time for the same MOP or MOE will also be pasted into different spreadsheets of the same workbook.

As you work with each MOP or MOE, you will want to save a new version of the workbook using "Save As," to keep a clean original version of each workbook for use with other MOPs and MOEs.

## Analyzing Data

In this section, we first discuss different types of variables and the descriptive statistics that correspond to them. We then describe different types of analysis that you are likely to use in your evaluation, and how to use the RAND Tool to implement them, with a focus on analysis of individual-level data. The RAND Tool is composed of seven workbooks in Microsoft Excel for different types of variables and analyses, as depicted in Figure 6.8 and referenced throughout this section. Finally, we discuss ways of modifying or expanding the RAND Tool to fit your PSC's needs, including how to use it for aggregated data instead of individual-level data.

### Descriptive Statistics for Different Variable Types

There are two types of variables likely to be relevant to your PSC's MOPs and MOEs. These types are continuous variables and dichotomous variables. These variables can be relevant at the level of individual participants, at the level of specific PSC sessions, or at the level of the

**Figure 6.8**  
**Schematic of Workbooks in RAND Tool Based on Variable Type and Analysis**

|                      | Continuous Variables               | Dichotomous Variables             |
|----------------------|------------------------------------|-----------------------------------|
| Comparison Over Time | 1_Tool_Continuous_Over_Time.xlsx   | 2_Tool_Dichotomous_Over_Time.xlsx |
| Pre/Post Comparison  | 3_Tool_Continuous_Pre_Post.xlsx    | 4_Tool_Dichotomous_Pre_Post.xlsx  |
| Comparison By Groups | 5_Tool_Continuous_By_Group.xlsx    | 6_Tool_Dichotomous_By_Group.xlsx  |
| Correlation          | 7_Tool_Continuous_Correlation.xlsx | –                                 |

RAND TL264-6.8

PSC as a whole. The way that basic descriptive statistics are calculated depends on the type of variable.

### ***Dichotomous Variables***

A dichotomous variable has only two possible outcomes (e.g., male or female, satisfied or not, has a resume or not), and these values are frequently represented as values of “0” and “1.” The relevant descriptive statistic for this type of variable is a proportion. For example, for FMEAP, one suggested MOE is the percentage of family members who have a completed resume. To calculate a percentage for dichotomous variables, you simply divide the numerator (e.g., number of family members receiving career coaching who have a completed resume) by the denominator (e.g., number of family members receiving career coaching who need a resume) and multiply by 100:

$$\frac{\text{numerator}}{\text{denominator}} \times 100 = \%$$

$$\frac{5 \text{ family members have a completed resume}}{16 \text{ family members need a resume}} = 0.313 \times 100 = 31.3\%$$

### **RAND Tool**

For dichotomous variables collected at the individual level, copy the columns representing the participant ID and the dichotomous variable and paste them into the relevant Excel workbook: *2\_Tool\_Dichotomous\_Over\_Time.xlsx* (Workbook 2), *4\_Tool\_Dichotomous\_Pre\_Post.xlsx* (Workbook 4), and *6\_Tool\_Dichotomous\_By\_Group.xlsx* (Workbook 6). Enter the codes for each outcome (e.g., 0, 1) and the corresponding description of each code (e.g., no, yes) in the spreadsheet. The workbooks will calculate the percentage indicated (e.g., “percent yes”). If you are tracking a dichotomous indicator over time, then paste the ID and dichotomous variables for the first time point in the Time 1 spreadsheet in *2\_Tool\_Dichotomous\_Over\_Time.xlsx* (Figure 6.9).

In this example, the dichotomous variable is coded 0 = “no” and 1 = “yes,” and the tool is configured to calculate the percentage of participants indicating “yes” on the dichotomous variable.

*Proportions Based on Aggregated Data.* Calculating the percentage of participants who have a resume assumes that response values are collected at the individual level (e.g., from each

**Figure 6.9**  
**Entering and Coding Dichotomous Variables in RAND Tool Workbook**

|   | A    | B                      | C | D                            | E | F | G      | H     | I   |
|---|------|------------------------|---|------------------------------|---|---|--------|-------|-----|
| 1 | ID   | Paste Time 1 data here |   |                              |   |   |        | No    | Yes |
| 2 | 2001 | 0                      |   | Enter codes for the variable |   |   |        | 0     | 1   |
| 3 | 2002 | 0                      |   |                              |   |   |        |       |     |
| 4 | 2003 | 0                      |   | Time Period of Data          |   |   | Jan-17 |       |     |
| 5 | 2004 | 0                      |   |                              |   |   |        |       |     |
| 6 | 2005 | 0                      |   | Percent: Yes                 |   |   |        | 31.3% |     |
| 7 | 2006 | 0                      |   |                              |   |   |        |       |     |

RAND TL264-6.9

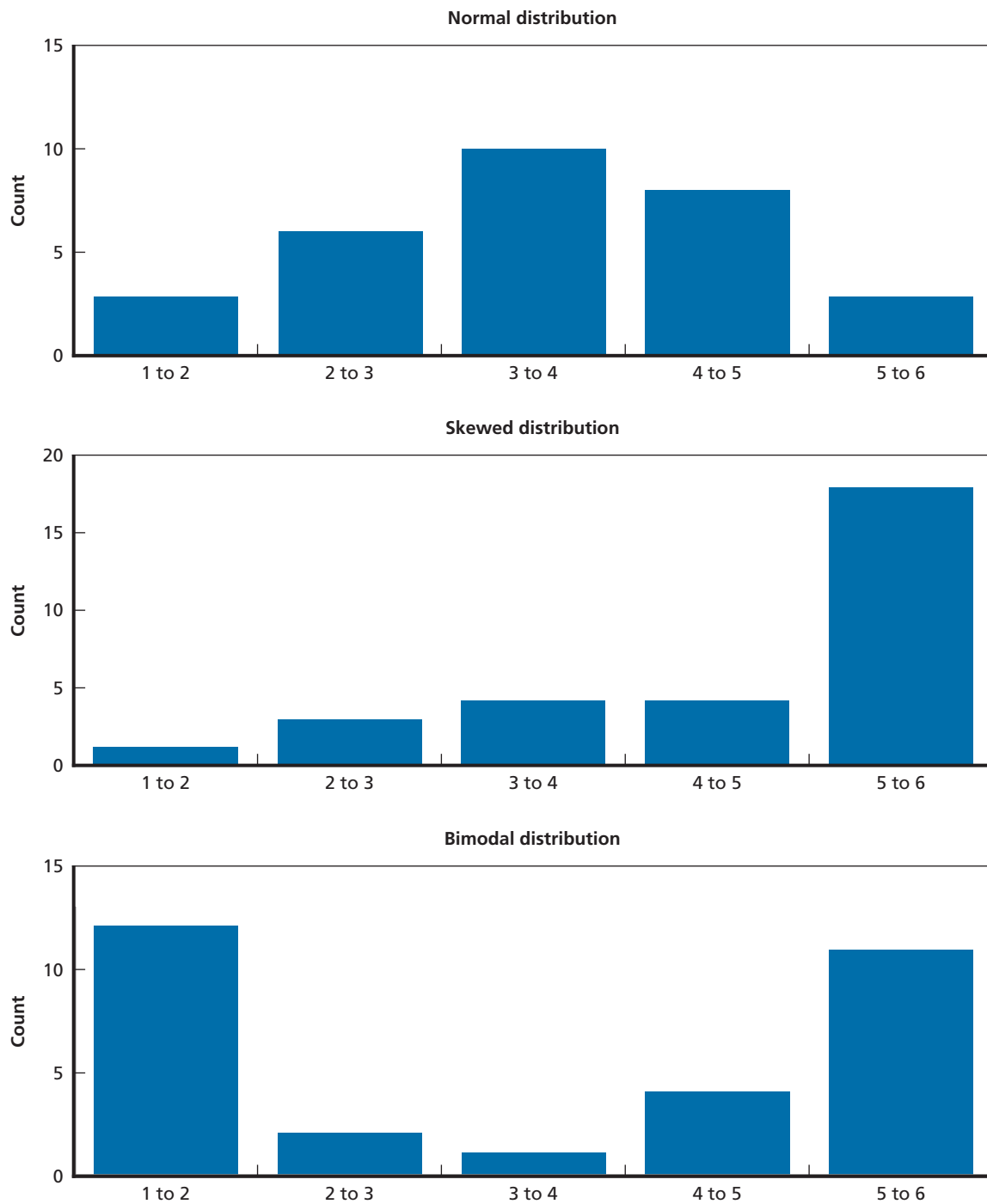
FMEAP participant who needs a resume). In other cases, an MOP or MOE may be a proportion or ratio formed from aggregated data, such as the proportion of required equipment pieces that are available (a suggested MOP for Outdoor Recreation Equipment Rental). In this case, the proportion would be calculated simply by dividing one numeric value (number of available equipment pieces) by another numeric value (number of required equipment pieces), instead of by recording values for different individuals to create a dichotomous variable.

### **Continuous Variables**

A continuous variable has multiple possible outcomes along a continuum. At the individual level, these values may be, for example, the number of job applications submitted by a participant or the percentage of items correct on a knowledge quiz. These values may also be a response scale (e.g., “very dissatisfied” to “very satisfied,” coded 1 to 5) for a single question or a published measure (e.g., the individual’s average score for a set of items measuring marital satisfaction). Or, at the aggregate level, you may be working with numeric values such as the number of individuals, sessions, or incidents. Sometimes continuous variables are collapsed into dichotomous variables; for example, when you want to examine the percentage of PSC participants who said they were “satisfied” or “very satisfied” with the PSC, you would calculate a percentage based on those responses versus the other responses on the scale (described further later). Usually, though, you want to examine all of the values along the continuum, which for individual-level data are best described using indicators of central tendency, like the average score or number, and indicators of distribution or variability of scores among data points.

Notably, this toolkit provides analyses used when continuous variables are normally distributed, meaning that they take the shape of a *bell curve* when the count of how many respondents selected each possible response is plotted in a histogram (i.e., a chart that shows the number of participants who responded with a given value). In Figure 6.10, we illustrate how a normal distribution looks, with respondents clustered at a small number of frequent response values, and a “tail” on both the left and right for response values less frequently selected. For comparison, we also show a skewed distribution (with most participants clustered at one point and a long tail on one side only) and a bimodal distribution (with participants clustered at two different points).

**Figure 6.10**  
**Sample Histograms of Continuous Variables with Differing Distributions**



NOTE: The horizontal X-axis in each distribution depicts values on a response scale from 1 to 5, and the vertical Y-axis in each depicts the number of individuals who selected each response value.

RAND TL264-6.10

A histogram can be generated for your data using an Excel add-on. It is preferable that your continuous variables be normally distributed, because this is the assumption on which the statistical tests in the RAND Tool rest. If you have reason to believe that a variable is not normally distributed, it is a good idea to generate a histogram and consult a statistician about the appropriate statistical test to use.

### RAND Tool

For individual-level continuous variables, copy the columns representing the participant ID and the continuous variable and paste them into the relevant Excel workbook: *1\_Tool\_Continuous\_Over\_Time.xlsx* (Workbook 1), *3\_Tool\_Continuous\_Pre\_Post.xlsx* (Workbook 3), and *5\_Tool\_Continuous\_By\_Group.xlsx* (Workbook 5). For scales, paste in the column containing each individual's average score for the scale items, described earlier—do not paste in the individual variables for the scale items. If you are tracking a continuous indicator over time, then paste the ID variable and continuous variable for the first time point in the Time 1 spreadsheet in *1\_Tool\_Continuous\_Over\_Time.xlsx* (Figure 6.11):

In this example, the individual averages calculated from a scale were being tracked over time, so the data were first sorted by ID then pasted into the spreadsheet. The tool automatically calculates the overall *average* of scores as the measure of central tendency. An alternative measure of central tendency that the tool also calculates is the *median*, the value at the middle of the distribution of the data (i.e., 50 percent of the data values are greater than the median and 50 percent are lower). If there are extreme values in your data set that are set apart from other values, it can be more helpful to look at the median than the average. For example, if one participant in FMEAP had 15 job applications but most others had about 5, the average value would be pulled upward by the extreme value of 15, whereas the median would not.

**Figure 6.11**  
Pasting Continuous Indicators into RAND Tool Workbook

| B1 |      | Paste Time 1 continuous data here |   |                     |               |        |  |
|----|------|-----------------------------------|---|---------------------|---------------|--------|--|
|    | A    | B                                 | C | D                   | E             | F      |  |
|    | ID   | Paste Time 1 continuous data here |   |                     |               |        |  |
| 1  |      |                                   |   |                     |               |        |  |
| 2  | 2001 | 2                                 |   | Time Period of Data |               | Jan-17 |  |
| 3  | 2002 | 3                                 |   |                     |               |        |  |
| 4  | 2003 | 4                                 |   |                     | Average:      | 2.9    |  |
| 5  | 2004 | 5                                 |   |                     | Median:       | 3.0    |  |
| 6  | 2005 | 4                                 |   | Standard Deviation: |               | 1.5    |  |
| 7  | 2006 | 5                                 |   |                     |               |        |  |
| 8  | 2007 | 4                                 |   |                     | Lower 95% CI: | 2.1    |  |
| 9  | 2008 | 5                                 |   |                     | Upper 95% CI: | 3.7    |  |
| 10 | 2009 | 1                                 |   |                     |               |        |  |
| 11 | 2010 | 3                                 |   |                     |               |        |  |

NOTE: If you are tracking individuals by their ID over time, sort your data by ID before pasting it into the spreadsheet. This will ensure that the data are analyzed correctly.

There are two indicators of distribution/variability among individuals on this measure: the *standard deviation*, an indicator of how widely values are spread around the overall average, and the upper and lower bounds, which contain 95 percent of the data values (also known as the *95 percent confidence interval*, or CI).

### Converting a Continuous Variable to a Dichotomous Variable

As mentioned previously, in some cases you may want to convert a continuous variable to a dichotomous variable. For example, if an MOP is the proportion of participants “satisfied” or “very satisfied” with the PSC, and participants were asked to rate their satisfaction on a continuous scale from 1 to 5, you would want to know the percentage of participants who had selected the response values for “satisfied” or “very satisfied” in their rating.

Consider an example where 4 means “satisfied” and 5 means “very satisfied.” In Figure 6.12, we show that you would add a new column in Excel for the dichotomous variable and use the IF() formula to tell each cell what values to print depending on the corresponding value in the column for the continuous variable.

In the IF() formula, you should first indicate a statement, which Excel refers to as a logical test (e.g., the statement that the response value for the continuous variable should be greater than 3).

After a comma, indicate the value to print if that logical test is true. In this example, because we are interested in the response values greater than 3 (i.e., respondents who were satisfied or very satisfied), we say that the value in the new column should be 1 for values that make the logical test true.

**Figure 6.12**  
Sample Use of Formula to Count Number of Times a Value Appears

| ID  | Satisfaction rating (continuous)       | Satisfied or very satisfied (dichotomous) |
|-----|----------------------------------------|-------------------------------------------|
| 201 | 2                                      | 0                                         |
|     | 1 = Very Dissatisfied                  | 0 = No (rating of 1, 2, or 3)             |
|     | 2 = Dissatisfied                       | 1 = Yes (rating of 4 or 5)                |
|     | 3 = Neither Satisfied Nor Dissatisfied |                                           |
|     | 4 = Satisfied                          |                                           |
|     | 5 = Very Satisfied                     |                                           |
| 213 | 2                                      | 0                                         |
| 214 | 3                                      | 0                                         |
| 215 | 1                                      | 0                                         |
| 216 | 1                                      | 0                                         |
| 301 | 3                                      | 0                                         |
| 302 | 4                                      | 1                                         |
| 303 | 4                                      | 1                                         |
| 304 | 5                                      | =IF(B15>3, 1, 0)                          |
| 305 | 4                                      | 1                                         |
| 306 | 5                                      | 1                                         |
| 307 | 3                                      | 0                                         |

After a second comma, indicate the value to print if the logical test is false. In this example, for response values that are 3 or lower (i.e., respondents who were only very dissatisfied, dissatisfied, or neither satisfied nor dissatisfied), we say that the value in the new column should be 0 for values that make the logical test false.

**Formula for converting a continuous variable to a dichotomous variable in a new column:**

`=IF(logical test, value to print if logical test is TRUE, value to print if logical test is FALSE)`

Example

`=IF(value for continuous variable > 3, 1, 0)`

Using this formula in this way assumes that all response values do fall within the expected range of 1 through 5. The formula in the new column can be dragged down from the top cell to apply to the full column. The new dichotomous variable can then be analyzed using the tools we provide for dichotomous variables.

### **Other Types of Variables**

Not all variables are dichotomous or continuous. Some variables have more than two values that do not lie along a distinct continuum (e.g., yes, no, don't know; associate degree, some college, bachelor's degree). If you find that you need to use variables in your evaluation that are not explicitly dichotomous or continuous, or are unsure of the type of variable you have, then consult a statistical expert for how to conduct the analysis.

### **Analysis Types**

Just as there are both continuous and dichotomous variables, there are multiple types of analysis relevant to different MOPs and MOEs. The primary three types that we discuss are continuous comparison over time, pre/post comparison, and comparison between groups. Each of these analyses can be done with either continuous or dichotomous variables, as shown in the RAND Tool. Another potentially useful analytic step that we discuss is assessing correlations between variables.

Throughout, we discuss the use of testing for *statistical significance* to see how strong the evidence is that a difference between two values is meaningful (rather than happening by chance). Tests for statistical significance yield what is known as a *p value*. When the *p value* is lower than a given threshold, then the comparison is interpreted as statistically significant and as stronger evidence of a difference. While the exact threshold is arbitrary, here we use the conventional cutoff of  $p < 0.05$ , corresponding to a confidence level of 0.95 or significance level of 5 percent.

### **Continuous Comparison over Time**

When tracking an MOP or MOE over time, there are situations where you want to monitor the variable at set intervals to see whether it generally increases or decreases over time (e.g., does average satisfaction with your PSC increase over consecutive months?), or you want to see whether a variable changes from one point to another (e.g., changes in participant attitudes from before taking a course to after taking the course). We first discuss the ongoing tracking of variables and the RAND tools used for tracking variables over time, and then discuss the analysis of pre- and postevent variables and RAND tools for those analyses.

### Ongoing Tracking of Indicators

Some MOPs and MOEs should be tracked over time to examine ongoing trends in the data. For example, if participation in the PSC is an indicator, then you will want to track whether participation increases, declines, or stays the same over time. If satisfaction with the PSC is an indicator, then you will want to track average satisfaction scores and/or percentage satisfied/very satisfied over time. The choice of time period to monitor MOPs/MOE is dependent on the particular indicator but could be weekly, monthly, quarterly, or annually.

### RAND Tool

For *dichotomous* variables collected at the individual level, the *2\_Tool\_Dichotomous\_Over\_Time.xlsx* workbook provides descriptive and statistical analysis of indicators. Following the instructions foregoing for pasting data into spreadsheets for descriptive statistics, data from the first time point would go into the “Time 1” spreadsheet, data from the second time point would go into the “Time 2” spreadsheet, etc. Enter the time period for the data from each spreadsheet (Figure 6.13):

The tool allows for four time periods of data, although trends in the data can be examined with fewer than four time periods.

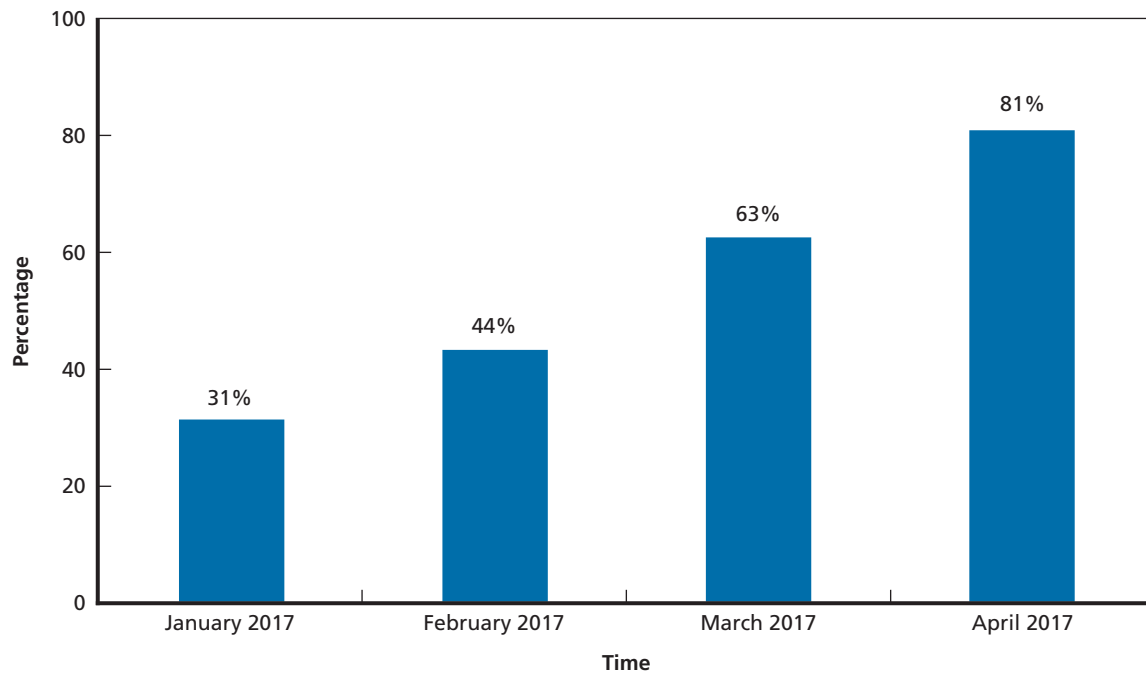
The final spreadsheet, labeled “Over Time Results,” displays a graph of the data over time and the results of statistical tests analyzing whether there was a significant difference between individual time points. The tests of differences between consecutive time points are shown first, followed by tests of differences between nonconsecutive time points (e.g., between month 1 and month 3) (Figure 6.14).

The statistical results test the probability that the percentages for two time points would occur based on chance alone, using a chi-square test for independence. The chi-square test examines the cross between the number of observations (i.e., response values for different participants or families) in each dichotomous category at each time point. It tests for whether these numbers differ from what would be expected if the pattern of responses were the same across time points. The criterion for determining a statistically significant difference is set at less than  $p < 0.05$ . In the tool, the  $p$ -value of the test is displayed, along with text describing whether the difference is significant or not. Keep in mind that the significance test does not test for the direction of the change, so scores could significantly increase or decrease, and the

**Figure 6.13**  
Entering Data into a Time 1 Spreadsheet

|   | A    | B                      | C | D                            | E      | F | G | H     | I   |
|---|------|------------------------|---|------------------------------|--------|---|---|-------|-----|
| 1 | ID   | Paste Time 1 data here |   |                              |        |   |   | No    | Yes |
| 2 | 2001 | 0                      |   | Enter codes for the variable |        |   |   | 0     | 1   |
| 3 | 2002 | 0                      |   |                              |        |   |   |       |     |
| 4 | 2003 | 0                      |   | Time Period of Data          | Jan-17 |   |   |       |     |
| 5 | 2004 | 0                      |   |                              |        |   |   |       |     |
| 6 | 2005 | 0                      |   | Percent: Yes                 |        |   |   | 31.3% |     |
| 7 | 2006 | 0                      |   |                              |        |   |   |       |     |

**Figure 6.14**  
**Sample Graph of a Dichotomous Indicator over Time**



RAND TL264-6.14

$p$ -value would not distinguish this difference. You would need to look at the accompanying percentages and graph to determine whether the change was an increase or decrease between time points. Figure 6.15 is an example using the previous figure.

In this example, data were collected monthly. The first box tests for significant differences in percentages across consecutive months (i.e., from January to February, February to March, and March to April), and the second box tests for significant differences between nonconsecutive months. Both tests are useful: The first set of tests shows whether there were significant incremental changes over each time period, and the second set of tests shows whether there were significant changes between periods of time that were further apart (e.g., from January to April). As you can see, none of the changes over consecutive months was significant at  $p < 0.05$ , but there were significant changes from January to April and from February to April. Looking at the graph, the changes between consecutive months were smaller than the longer-term changes, so it makes sense that the tests for consecutive months were not significant, but two of the tests for longer periods of time were significant.

For *continuous* variables collected at the individual level, the *1\_Tool\_Continuous\_Over\_Time.xlsx* workbook provides descriptive and statistical analysis of indicators. Following the instructions previous for pasting data into spreadsheets for descriptive statistics, you would enter data from the first time point into the “Time 1” spreadsheet, data from the second time point into the “Time 2” spreadsheet, etc. Enter the time period for the data from each spreadsheet. The tool allows for four time periods of data, although trends in the data can be examined with fewer than four time periods.

The final spreadsheet, labeled “Over Time Results,” displays a graph of the averages over time and the results of statistical tests analyzing whether there was a significant difference

**Figure 6.15**  
**Sample Results of the Dichotomous**  
**Indicator Shown in Figure 6.14**

| Change over Time: Differences in Percentages over Consecutive Times        |      |                 |  |  |
|----------------------------------------------------------------------------|------|-----------------|--|--|
| Jan-17                                                                     | to   | Feb-17          |  |  |
| p =                                                                        | 0.47 | Not Significant |  |  |
| Feb-17                                                                     | to   | Mar-17          |  |  |
| p =                                                                        | 0.29 | Not Significant |  |  |
| Mar-17                                                                     | to   | Apr-17          |  |  |
| p =                                                                        | 0.24 | Not Significant |  |  |
| Change over Time: Differences in Percentages Between Non-Consecutive Times |      |                 |  |  |
| Jan-17                                                                     | to   | Mar-17          |  |  |
| p =                                                                        | 0.22 | Not Significant |  |  |
| Jan-17                                                                     | to   | Apr-17          |  |  |
| p =                                                                        | 0.00 | Significant     |  |  |
| Feb-17                                                                     | to   | Apr-17          |  |  |
| p =                                                                        | 0.03 | Significant     |  |  |

RAND TL264-6.15

between individual time points. The tests of differences between consecutive time points are shown first, followed by tests of differences between nonconsecutive time points (e.g., between month 1 and month 3).

*IMPORTANT: If you have tracked the same individuals over time and have sorted the data for each time point by ID, write “yes” in the yellow cell responding to the question “Are the same individuals tracked over time by ID?” If you did not track the same individuals over time by their ID, write “no” in this cell. If you write “yes” and the data are not sorted by ID or one or more individual IDs do not match across time, a warning will appear: “ERROR: at least one ID does not match.” If you receive this warning, re-sort and paste the data or identify the nonmatching ID number(s) and correct the error. See Figure 6.16.*

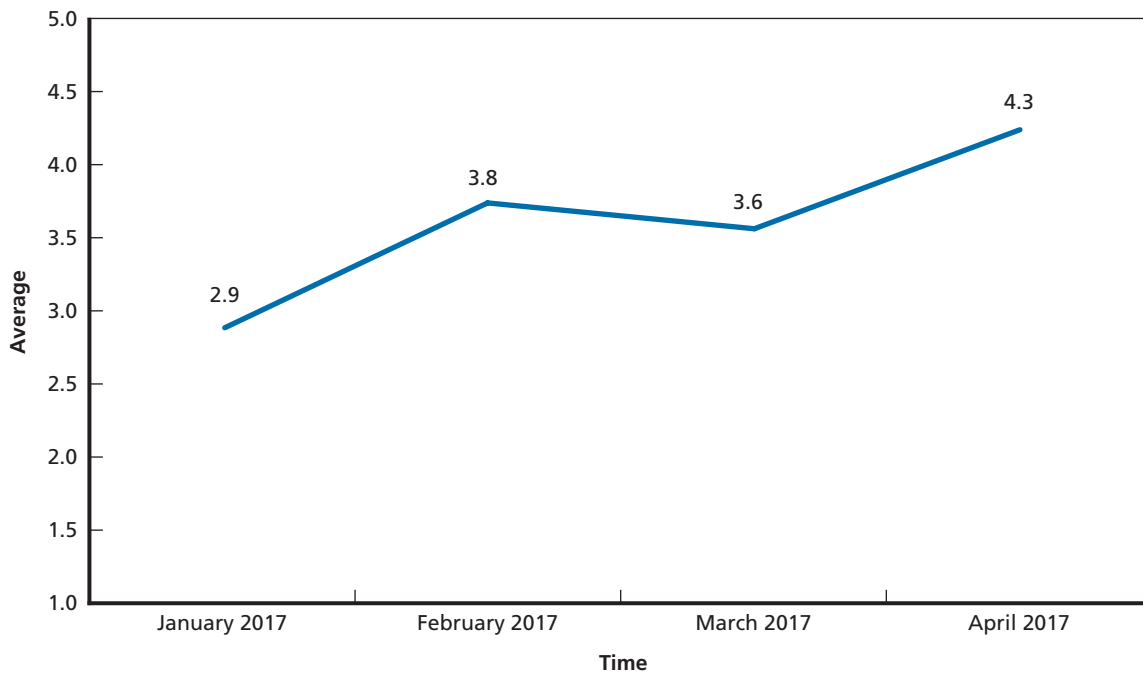
Figure 6.17 is a sample graph of a continuous indicator over time.

**Figure 6.16**  
**Write “Yes” or “No” in Yellow Cell for Tracking of Individuals**

|                       |        |        |        |   |   |                                                   |   |   |   |   |   |   |   |   |    |
|-----------------------|--------|--------|--------|---|---|---------------------------------------------------|---|---|---|---|---|---|---|---|----|
| O1                    |        |        |        |   |   |                                                   |   |   |   |   |   |   |   |   |    |
| fx no                 |        |        |        |   |   |                                                   |   |   |   |   |   |   |   |   |    |
| A                     | B      | C      | D      | E | F | G                                                 | H | I | J | K | L | M | N | O |    |
| Average Satisfaction: |        |        |        |   |   | Are the same individuals tracked over time by ID? |   |   |   |   |   |   |   |   | no |
| Jan-17                | Feb-17 | Mar-17 | Apr-17 |   |   |                                                   |   |   |   |   |   |   |   |   |    |
| 2.9                   | 3.8    | 3.6    | 4.3    |   |   |                                                   |   |   |   |   |   |   |   |   |    |

RAND TL264-6.16

**Figure 6.17**  
**Sample Graph of a Continuous Indicator over Time**



RAND TL264-6.17

The statistical results test the probability that the means for two time points are significantly different from each other, using a t-test. The t-test test examines the averages at the two time points, relative to the variation in data around the averages, and determines whether the two averages are significantly different from one another. The statistical formula is different when the same individuals are tracked over time by ID, so it is very important to indicate whether you tracked individuals by ID or not. The criterion for determining a statistically significant difference is set at less than  $p < 0.05$ . In the tool, the  $p$ -value of the test is displayed, along with text describing whether the difference is significant or not. Keep in mind that the significance test does not test for the direction of the change, so scores could significantly increase or decrease, and the  $p$ -value would not distinguish this difference. You would need to look at the accompanying percentages and graph to determine whether the change was an increase or decrease between time points.

Sample results of the continuous indicator shown in Figure 6.17 are shown in Figure 6.18.

In this example, data were collected monthly. The first box shows the results of tests for significant differences in averages across consecutive months (i.e., from January to February, February to March, and March to April), and the second box shows the results of tests for significant differences between nonconsecutive months. Both tests are useful: The first tests show whether there were significant incremental changes over each time period, and the second tests show whether there were significant changes between periods of time that were further apart (e.g., from January to April). As you can see, the change from March to April was significant, and none of the other changes over consecutive months was significant at  $p < 0.05$ . There was also a significant change from January to April. Looking at the graph, the average increased

**Figure 6.18**  
**Sample Results of the Continuous Indicator**  
**Shown in Figure 6.17**

| Change over Time: Average Ratings<br>over Consecutive Times        |      |                 |  |  |
|--------------------------------------------------------------------|------|-----------------|--|--|
| Jan-17                                                             | to   | Feb-17          |  |  |
| p =                                                                | 0.06 | Not Significant |  |  |
| Feb-17                                                             | to   | Mar-17          |  |  |
| p =                                                                | 0.57 | Not Significant |  |  |
| Mar-17                                                             | to   | Apr-17          |  |  |
| p =                                                                | 0.01 | Significant     |  |  |
|                                                                    |      |                 |  |  |
|                                                                    |      |                 |  |  |
|                                                                    |      |                 |  |  |
| Change over Time: Average Ratings<br>Between Non-Consecutive Times |      |                 |  |  |
| Jan-17                                                             | to   | Mar-17          |  |  |
| p =                                                                | 0.12 | Not Significant |  |  |
| Jan-17                                                             | to   | Apr-17          |  |  |
| p =                                                                | 0.00 | Significant     |  |  |
| Feb-17                                                             | to   | Apr-17          |  |  |
| p =                                                                | 0.11 | Not Significant |  |  |

RAND TL264-6.18

from January to February, leveled off from February to March, and then increased again from March to April.

*Considerations and Cautions.* Tracking ongoing trends in MOPs and MOEs is an effective way to monitor how your PSC is functioning over time. Favorable trends—increases in positive indicators (e.g., knowledge), decreases in negative indicators (e.g., problem behaviors), or steady trends at acceptable levels (e.g., consistent high attendance)—suggest that the PSC is serving its intended function with the targeted population. However, it is important to keep in mind that favorable trends do not prove that the PSC caused the trend: it could be that some other factor caused the favorable trend (e.g., people could have improved on their own, even without the PSC; they might have used another PSC that helped them with their problem). We discuss potential sources of misinterpretation in greater detail at the end of the chapter.

### Pre/Post Comparison

Statistical comparisons of pre- and postevent data are similar to those used to compare time points from the ongoing tracking of indicators. The main differences are that you have only two time points (e.g., before training and after using the PSC), and individual-level data are almost always tracked by individual ID numbers.

For **dichotomous** indicators collected at the individual level, follow the instructions foregoing for ongoing tracking of dichotomous indicators, except use the *4\_Tool\_Dichotomous\_Pre\_Post.xlsx* workbook to enter the pre and post dichotomous data. The pre data spreadsheet is labeled *Pre-PSC Scores*, and the post data spreadsheet is labeled *Post-PSC Scores*.

The Pre to Post PSC Changes spreadsheet contains a graph of the pre to post percentages and the results of the chi-square test of independence, evaluated at  $p < 0.05$ . For example, results of a pre to post change in a dichotomous indicator might be illustrated by Figure 6.19:

In this example, pre-PSC scores on the dichotomous variable were at 19 percent, and increased to 75 percent after the PSC. This change was significant at  $p < 0.05$  according to the chi-square test.

For **continuous** indicators collected at the individual level, follow the instructions preceding for ongoing tracking of continuous indicators, except use the *3\_Tool\_Continuous\_Pre\_Post.xlsx* workbook to enter the pre and post continuous data. The pre data spreadsheet is labeled *Pre-PSC Scores*, and the post data spreadsheet is labeled *Post-PSC Scores*.

The Pre to Post PSC Changes spreadsheet contains a graph of the pre to post averages and the results of the t-test comparing pre and post averages, evaluated at  $p < 0.05$ .

*IMPORTANT: If you have tracked the same individuals pre- and post-PSC and have sorted the data for each time point by ID, write “yes” in the yellow cell responding to the question “Are the same individuals tracked over time by ID?” If you did not track the same individuals pre- and post-PSC by their ID, write “no” in this cell. If you write “yes” and the data are not sorted by ID or one or more individual IDs do not match across time, a warning will appear: “ERROR: at least one ID does not match.” If you receive this warning, re-sort and paste the data or identify the nonmatching ID number(s) and correct the error.*

For example, Figure 6.20 shows what results of a pre to post change in a continuous indicator might be:

Note that the yellow cell indicating whether the same individuals are tracked over time is entered as “yes,” which adapts the t-test to account for tracking. In this example, pre-PSC average scores on the continuous variable were at 2.38 and increased to 4.06 after the PSC. This change was significant at  $p < 0.05$ , according to the t-test.

*Considerations and Cautions.* Examining pre- to post-PSC changes in MOPs and MOEs is an important tool for understanding the potential impact the PSC is having on participants.

**Figure 6.19**  
Results of a Pre to Post Change in a Dichotomous Indicator

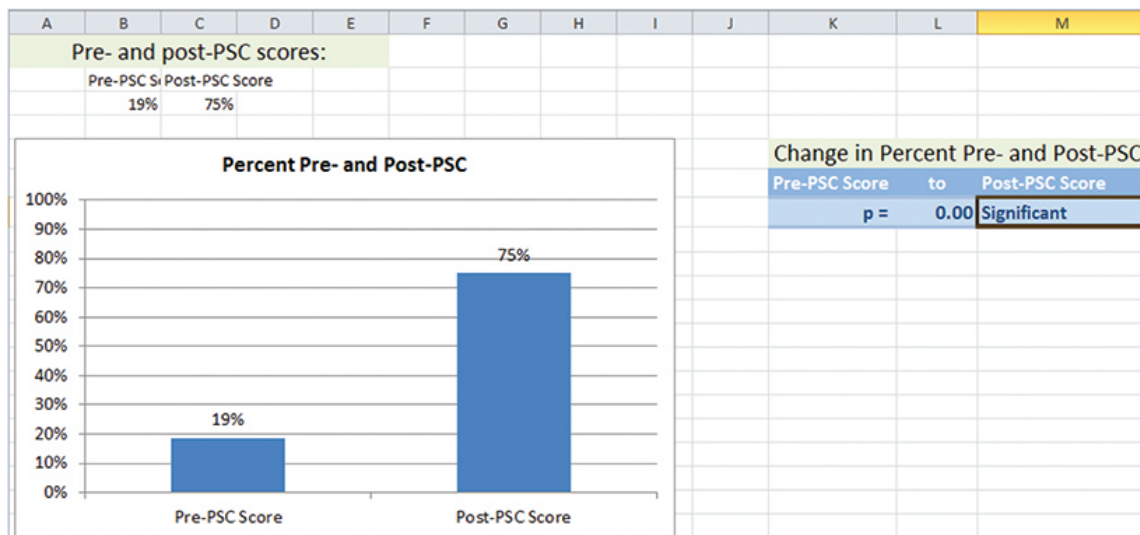
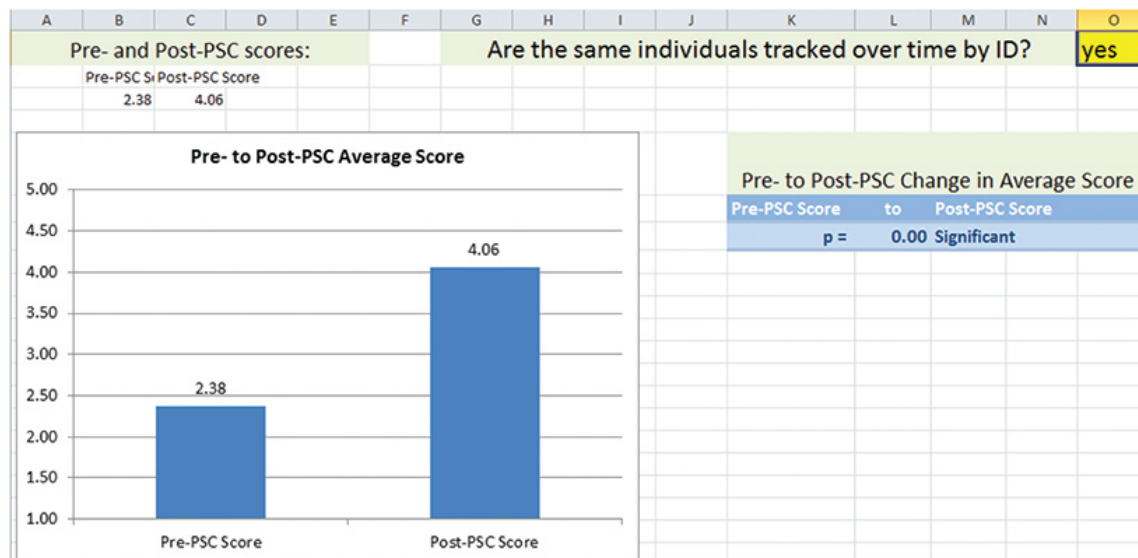


**Figure 6.20**  
**Notional Results of a Pre to Post Change in a Continuous Indicator**



RAND TL264-6.20

Favorable changes in indicators suggest that the PSC is serving its intended function with the targeted population. However, as with examining trends over time, it is important to keep in mind that favorable changes pre- to post-PSC do not prove that the PSC caused the change: It could be that some other factor caused the favorable change (e.g., people could have improved on their own, even without the PSC; they might have used another PSC that helped them with their problem). We discuss potential sources of misinterpretation in greater detail at the end of the chapter.

### **Comparison Between Groups**

There are times when you want to compare MOPs or MOEs between groups of people. For example, you might want to compare community connectedness between people who have participated in Outdoor Recreation Equipment Rental and those who have not participated.

#### **RAND Tool**

For **dichotomous** variables collected at the individual level, the *6\_Tool\_Dichotomous\_By\_Group.xlsx* workbook provides the statistical analysis of differences between groups. Following the instructions preceding for pasting data into spreadsheets for descriptive statistics, you would enter data from the first group (e.g., PSC users) into the “Group 1” spreadsheet, and data from the second group (e.g., PSC nonusers) would go into the “Group 2” spreadsheet. Make sure that the group codes entered in the spreadsheet match the appropriate group category (e.g., that PSC users are all coded as “1” and that data only from individuals with a 1 code are entered into the spreadsheet for users), and enter the name of the group in each spreadsheet, along with the codes for the dichotomous values (e.g., 0 and 1 for “No” and “Yes”) (Figure 6.21).

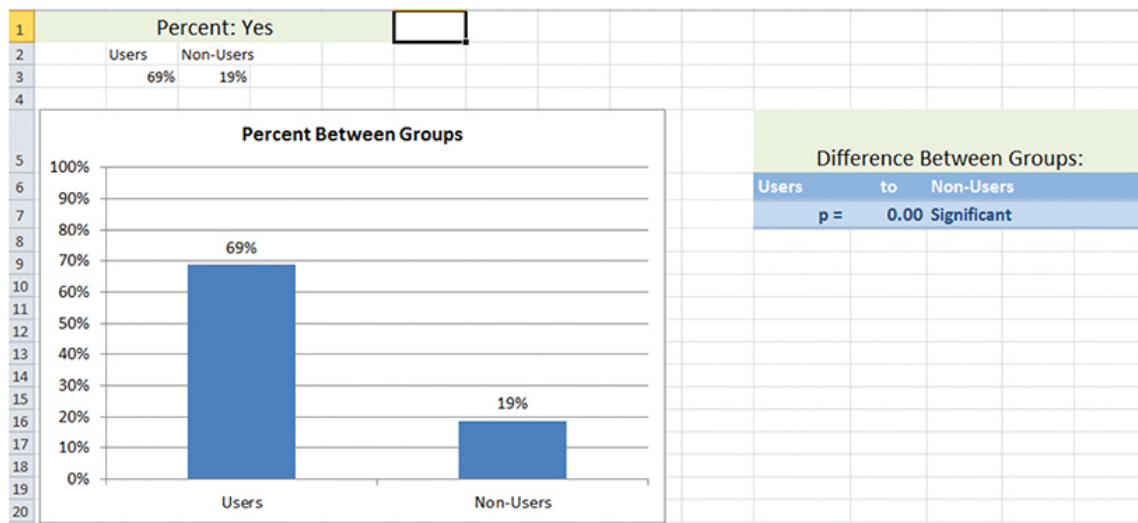
The final spreadsheet, labeled “Group Difference Results,” displays a graph of the data by group and the results of statistical tests analyzing whether there was a significant difference between the two groups (Figure 6.22).

**Figure 6.21**  
Entering and Coding Group Data in the RAND Tool Workbook

| H4    |      |                               |       |   |                              |   |   |       |       |     |
|-------|------|-------------------------------|-------|---|------------------------------|---|---|-------|-------|-----|
| Users |      |                               |       |   |                              |   |   |       |       |     |
|       | A    | B                             | C     | D | E                            | F | G | H     | I     | J   |
|       |      | Paste<br>Group 1<br>data here | Group |   |                              |   |   |       |       |     |
| 1     | ID   |                               |       |   |                              |   |   |       | No    | Yes |
| 2     | 2001 | 0                             | 1     |   | Enter codes for the variable |   |   |       | 0     | 1   |
| 3     | 2002 | 1                             | 1     |   |                              |   |   |       |       |     |
| 4     | 2003 | 1                             | 1     |   | Group name:                  |   |   | Users |       |     |
| 5     | 2004 | 1                             | 1     |   |                              |   |   |       |       |     |
| 6     | 2005 | 1                             | 1     |   | Percent: Yes                 |   |   |       | 68.8% |     |
| 7     | 2006 | 1                             | 1     |   |                              |   |   |       |       |     |

RAND TL264-6.21

**Figure 6.22**  
The Group Difference Results Spreadsheet



RAND TL264-6.22

In this example, PSC users scored 69 percent on the dichotomous indicator, and nonusers scored 19 percent. This difference was significant at  $p < 0.05$ , according to the chi-square test.

For **continuous** indicators collected at the individual level, the *5\_Tool\_Continuous\_By\_Group.xlsx* workbook provides the statistical analysis of differences between groups. Following the instructions foregoing for pasting data into spreadsheets for descriptive statistics, data from the first group (e.g., PSC users) would go into the “Group 1” spreadsheet, and data from the second group (e.g., PSC nonusers) would go into the “Group 2” spreadsheet. Make sure that the group codes entered in the spreadsheet match the appropriate group category (e.g., that PSC users are all coded as “1,” and that data only from individuals with a 1 code are entered into the spreadsheet for users), and enter the name of the group in each spreadsheet (Figure 6.23).

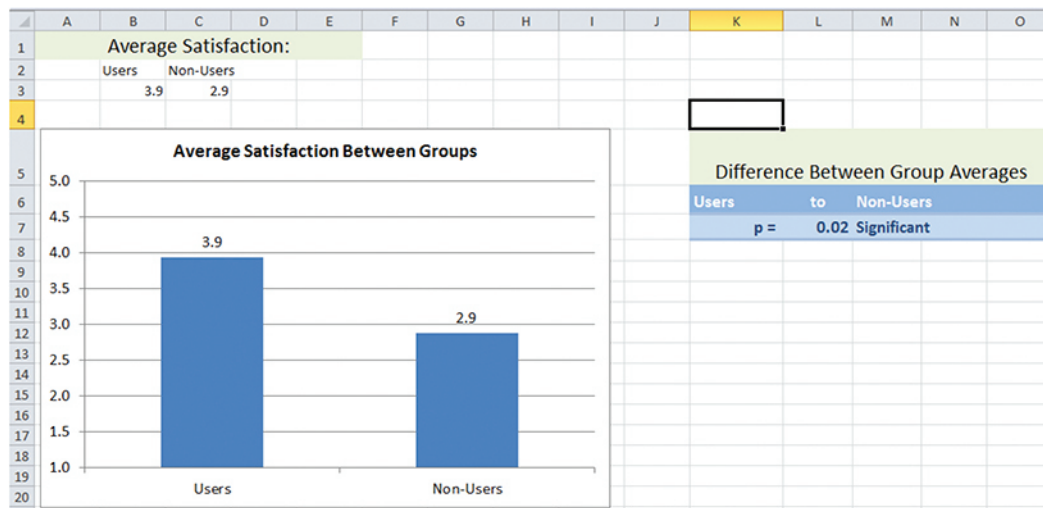
**Figure 6.23**  
**Entering Group Data into the RAND Tool Spreadsheet**

| G2       |     |                            |       |   |   |                     |       |
|----------|-----|----------------------------|-------|---|---|---------------------|-------|
| fx Users |     |                            |       |   |   |                     |       |
|          | A   | B                          | C     | D | E | F                   | G     |
| 1        | ID  | Paste Group<br>1 data here | Group |   |   |                     |       |
| 2        | 201 | 3                          | 1     |   |   | Group name:         | Users |
| 3        | 202 | 4                          | 1     |   |   |                     |       |
| 4        | 203 | 4                          | 1     |   |   | Average:            | 3.9   |
| 5        | 204 | 5                          | 1     |   |   | Median:             | 1.0   |
| 6        | 205 | 4                          | 1     |   |   | Standard Deviation: | 0.9   |
| 7        | 206 | 5                          | 1     |   |   |                     |       |
| 8        | 207 | 4                          | 1     |   |   | Lower 95% CI:       | 3.5   |
| 9        | 208 | 5                          | 1     |   |   | Upper 95% CI:       | 4.4   |
| 10       | 209 | 2                          | 1     |   |   |                     |       |

RAND TL264-6.23

The Group Difference Results spreadsheet displays a graph of the data by group and the results of statistical tests analyzing whether there was a significant difference between the two groups (Figure 6.24).

**Figure 6.24**  
**Group Difference Results Spreadsheet**



RAND TL264-6.24

In this example, average scores for users on the continuous variable were 3.9, and average scores for nonusers were 2.9. This difference was significant at  $p < 0.05$ , according to the t-test.

*Considerations and Cautions.* Examining differences between groups on MOPs and MOEs is another way to monitor how your PSC is affecting your target population. For example, finding that PSC users significantly differ from nonusers on key indicators (e.g., knowledge) suggests that the PSC had a positive impact on users. However, it could be that users were already

different from nonusers prior to seeking out the PSC (e.g., maybe those with more knowledge of family problems were more likely to seek out training on how to cope with family problems) or that users were somehow more inclined to process and understand the material than nonusers would have been (e.g., users who sought out training on family problems were more motivated to learn about family problems). Unless you randomly assign people to use or not use a PSC—thus ensuring that users did not have preexisting differences that might confound the impact of the PSC—you should use caution in interpreting group differences because they cannot be assumed to be causal. It could be that some other factor caused the favorable difference between groups. We discuss potential sources of misinterpretation in greater detail at the end of the chapter.

### ***Correlations Between Continuous Variables***

In addition to the analyses preceding, it can be useful to understand how two continuous variables are related to one another. For example, you might want to examine whether the number of Child and Youth Program activities that a family attends is related to a child having a lower number of disciplinary episodes, so you would conduct a correlation analysis to assess the strength of the relationship (higher correlation values indicate a stronger relationship).

Continuous variables are always tracked by individual ID numbers, so the data for the two variables must be entered with a participant ID. If the two variables are stored in separate spreadsheets, make sure to sort each spreadsheet by ID prior to pasting the variables into the RAND Tool, and that only individuals with data from both variables are included in the analysis tool.

Once you have confirmed that both variables contain only the data from the same participants, copy the columns representing the participant ID and the two continuous variables and paste them into the *7\_Tool\_Continuous\_Correlation.xlsx* workbook (Workbook 7). The “Correlation” cell will display the Pearson Correlation Coefficient between the two variables. The correlation coefficient varies from  $-1$  to  $1$ , with negative values indicating a negative correlation (i.e., higher scores on one variable are associated with lower scores on the other variable) and positive values indicating a positive correlation (i.e., higher scores on one variable are associated with higher scores on the other variable). Coefficients near  $0$  represent no correlation between variables (i.e., values for the two variables are not related to one another). Variables with correlation coefficients of about  $\pm 0.30$  are considered weakly correlated, correlation coefficients of about  $\pm 0.50$  are considered moderately correlated, and correlation coefficients of about  $\pm 0.70$  are considered strongly correlated. In Figure 6.25, the two variables are moderately correlated.

*Considerations and Cautions.* The correlation between two continuous variables can demonstrate important relationships among your MOPs and MOEs. For example, finding that there is a moderate or strong correlation between number of learning sessions attended by PSC users and tests of knowledge from the sessions suggests that the PSC sessions are helping users understand the material being taught. Of course, the saying that “correlation is not causation” is true here as well: It could be that users who attended more sessions were already more knowledgeable than those who attended fewer sessions, were more inclined to process and understand the material than those who attended fewer sessions, and/or were more motivated to learn the material than those who attended fewer sessions. As with the other tests reviewed in this section, you should use caution in interpreting correlation coefficients between variables

**Figure 6.25**  
**Correlation Workbook Showing Moderately Correlated Variables**

|   | A    | B                                              | C                                              | D           | E | F    | G |
|---|------|------------------------------------------------|------------------------------------------------|-------------|---|------|---|
|   |      | Paste<br>continuous<br>variable 1<br>data here | Paste<br>continuous<br>variable 2<br>data here |             |   |      |   |
| 1 | ID   |                                                |                                                |             |   |      |   |
| 2 | 2001 | 1                                              | 3                                              |             |   |      |   |
| 3 | 2002 | 2                                              | 3                                              |             |   |      |   |
| 4 | 2003 | 3                                              | 4                                              | Correlation |   | 0.56 |   |
| 5 | 2004 | 4                                              | 4                                              |             |   |      |   |

RAND TL264-6.25

because the relationship cannot be assumed to be causal. We discuss potential sources of misinterpretation in greater detail at the end of the chapter.

### Adapting the RAND Tool

The RAND Tool offers a starting point for the different analyses you may want to conduct. There are ways to adapt it for different uses depending on your PSC. For example, you may want to include more time points. However, once you reach a larger period of time for comparison, you can consider collapsing different time points. If you begin comparing across periods of one month, for example, you can transition to comparing across quarters, and ultimately comparing across years.

While the tool has been designed to be able to accommodate data collected from individual respondents, it can also be used to look at trends over time in aggregated data, more generally. We describe this use more next.

### Using the RAND Tool for Aggregated Data

Many MOPs (and some MOEs) take the form of a proportion at the aggregate level, such as number of attendees at a program out of number of spaces available for the program, or number of weekly hours of operation out of number of operating hours required each week. These proportions can be analyzed in the same way as a proportion derived from a dichotomous variable, with a different row dedicated to each “unit” (e.g., spaces in program, hours of operation). For example, to analyze the proportion of equipment pieces available out of 20 required equipment pieces, you would include 20 rows in the appropriate time point spreadsheet in the workbook for continuous comparison over time of dichotomous variables. You would fill in 1s for the number of rows representing available equipment pieces, and 0s for the number of rows representing unavailable equipment pieces. Thus, you would be entering “individual-level data,” but with the individual units being equipment pieces instead of participants or families. The workbook can then be used as usual.

Many MOPs also take the form of a simple numeric value, such as number of attendees, number of hours volunteered, or number of facilities issues per year. These values are not considered variables, though, as there is not a series of data points about different individuals or other units. While you will not be running statistical tests, you can visualize the numeric value over time, using an Excel chart of the same form as the one in the workbook for looking

at continuous variables over time. We caution against using the RAND Tool for this purpose, however, to avoid confusion about the statistical tests that it contains (which will not be applicable, as they are designed for individual-level data). With enough numeric values from different time points, gathered over a long period of time, a statistical test for a positive trend can be done. However, we suggest for now focusing on whether the numeric values are meeting a goal you may have in mind (discussed more next).

## Identifying Trends, Performance Outcomes, and Effectiveness

After conducting analysis of the data, the next step is to interpret the results and identify patterns that indicate PSC trends, performance outcomes, and effectiveness. In this section, we first provide a worksheet to assist in identifying and interpreting trends for different MOPs and MOEs. We then summarize potential sources of misinterpretation that should be kept in mind when interpreting the results of your analyses.

### Identifying and Interpreting Trends

Worksheet 6.1 is a blank worksheet to document your findings for different MOPs and MOEs and whether the PCS met, fell short of, or exceeded the desired outcomes. Example Worksheet 6.1 offers examples from L.I.N.K.S. MOP and MOE analyses that could take place.

To use the worksheet, start by documenting in the first column what your intended outcome was for each MOP and MOE analyzed. In the second and third columns, respectively, record what comparison you made in the above analysis section, and what the actual outcome of the comparison was. In the fourth column, “What is the trend?,” indicate whether the value of primary interest is better, the same, or worse than the comparison value(s). Depending on the type of comparison being made, the value of primary interest may be

- the most *recent* value (versus earlier values, for continuous comparison over time)
- the *postprogram* value for participants (versus the preprogram value, for pre/post comparison)
- the value for individuals who *participated* in the program (versus for individuals who did not participate, for comparison between groups)
- the *recorded* value (versus the metric, for comparison to a metric).

In the fifth column, record whether this trend met, missed (fell short of), or exceeded the desired outcome from the first column. And finally, if you missed your goal, indicate in the last column that action is needed. The next chapter will help in thinking about next steps if action is needed.

For continuous comparison over time, you will want to think about whether your goal was improving an MOP or MOE and by how much, or if you simply wanted to make sure that the value stays consistent and does not get worse over time. You can set a metric that is your goal value (for example, increasing satisfaction ratings from 3.8 to 4.5, out of 5), and may want to think about a reasonable goal for the amount of time it would take to reach that value (e.g., over the next year).

In some cases, you may simply be comparing a value to a metric that was your goal, instead of doing one of the analyses described in this chapter and in the RAND Tool. For

**Worksheet 6.1****Review Program Outcomes***Adapted from Acosta et al., 2013, and Hunter et al., 2015*

| Desired Outcome for MOP/ MOE (Value, Difference, or Change)                             | Comparison Conducted                                                                                                                                                                                                                         | Actual Outcome for MOP/MOE |                                                                                                                                 | Met Expectations?                                                                                                                                               | Action Needed?                                              |
|-----------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|---------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------|
| <input type="checkbox"/> MOP<br><input type="checkbox"/> MOE<br><br><b>Description:</b> | <input type="checkbox"/> Continuous comparison over time<br><input type="checkbox"/> Pre/post comparison<br><input type="checkbox"/> Comparison between groups<br><input type="checkbox"/> Comparison with metric<br><br><b>Description:</b> |                            | <b>What is the trend?</b><br><input type="checkbox"/> Better<br><input type="checkbox"/> Same<br><input type="checkbox"/> Worse | <b>Did this meet your expectations for the program?</b><br><input type="checkbox"/> Met<br><input type="checkbox"/> Missed<br><input type="checkbox"/> Exceeded | <input type="checkbox"/> Yes<br><input type="checkbox"/> No |
| <input type="checkbox"/> MOP<br><input type="checkbox"/> MOE<br><br><b>Description:</b> | <input type="checkbox"/> Continuous comparison over time<br><input type="checkbox"/> Pre/post comparison<br><input type="checkbox"/> Comparison between groups<br><input type="checkbox"/> Comparison with metric<br><br><b>Description:</b> |                            | <b>What is the trend?</b><br><input type="checkbox"/> Better<br><input type="checkbox"/> Same<br><input type="checkbox"/> Worse | <b>Did this meet your expectations for the program?</b><br><input type="checkbox"/> Met<br><input type="checkbox"/> Missed<br><input type="checkbox"/> Exceeded | <input type="checkbox"/> Yes<br><input type="checkbox"/> No |
| <input type="checkbox"/> MOP<br><input type="checkbox"/> MOE<br><br><b>Description:</b> | <input type="checkbox"/> Continuous comparison over time<br><input type="checkbox"/> Pre/post comparison<br><input type="checkbox"/> Comparison between groups<br><input type="checkbox"/> Comparison with metric<br><br><b>Description:</b> |                            | <b>What is the trend?</b><br><input type="checkbox"/> Better<br><input type="checkbox"/> Same<br><input type="checkbox"/> Worse | <b>Did this meet your expectations for the program?</b><br><input type="checkbox"/> Met<br><input type="checkbox"/> Missed<br><input type="checkbox"/> Exceeded | <input type="checkbox"/> Yes<br><input type="checkbox"/> No |
| <input type="checkbox"/> MOP<br><input type="checkbox"/> MOE<br><br><b>Description:</b> | <input type="checkbox"/> Continuous comparison over time<br><input type="checkbox"/> Pre/post comparison<br><input type="checkbox"/> Comparison between groups<br><input type="checkbox"/> Comparison with metric<br><br><b>Description:</b> |                            | <b>What is the trend?</b><br><input type="checkbox"/> Better<br><input type="checkbox"/> Same<br><input type="checkbox"/> Worse | <b>Did this meet your expectations for the program?</b><br><input type="checkbox"/> Met<br><input type="checkbox"/> Missed<br><input type="checkbox"/> Exceeded | <input type="checkbox"/> Yes<br><input type="checkbox"/> No |

**Example Review Program Outcomes Using L.I.N.K.S.***Adapted from Acosta et al., 2013, and Hunter et al., 2015*

| Desired Outcome for MOP/ MOE (Value, Difference, or Change)                                                                                                                                                                    | Comparison Conducted                                                                                                                                                                                                                                                                                                                                                                                                | Actual Outcome for MOP/MOE                                                            |                                                                                                                                            | Met Expectations?                                                                                                                                                          | Action Needed?                                                         |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------|
| <input checked="" type="checkbox"/> MOP<br><input type="checkbox"/> MOE<br><br><i>Description:</i><br>Participant satisfaction with L.I.N.K.S. should increase (or not decrease) over time                                     | <input checked="" type="checkbox"/> Continuous comparison over time<br><input type="checkbox"/> Pre/post comparison<br><input type="checkbox"/> Comparison between groups<br><input type="checkbox"/> Comparison with metric<br><br><i>Description:</i><br>Satisfaction ratings for L.I.N.K.S. military culture workshops were compared from 2017 to 2018                                                           | Average satisfaction was significantly lower in 2018 (vs. 2017)                       | <i>What is the trend?</i><br><input type="checkbox"/> Better<br><input type="checkbox"/> Same<br><input checked="" type="checkbox"/> Worse | <i>Did this meet your expectations for the program?</i><br><input type="checkbox"/> Met<br><input checked="" type="checkbox"/> Missed<br><input type="checkbox"/> Exceeded | <input checked="" type="checkbox"/> Yes<br><input type="checkbox"/> No |
| <input checked="" type="checkbox"/> MOP<br><input type="checkbox"/> MOE<br><br><i>Description:</i><br>100% of staff should receive required annual training that provides updates on program curricula                         | <input type="checkbox"/> Continuous comparison over time<br><input type="checkbox"/> Pre/post comparison<br><input type="checkbox"/> Comparison between groups<br><input checked="" type="checkbox"/> Comparison with metric<br><br><i>Description:</i><br>Proportion of staff who received required annual training in 2017 was compared with 100%                                                                 | Only 75% of L.I.N.K.S. trainers received required annual training, compared with 100% | <i>What is the trend?</i><br><input type="checkbox"/> Better<br><input type="checkbox"/> Same<br><input checked="" type="checkbox"/> Worse | <i>Did this meet your expectations for the program?</i><br><input type="checkbox"/> Met<br><input checked="" type="checkbox"/> Missed<br><input type="checkbox"/> Exceeded | <input checked="" type="checkbox"/> Yes<br><input type="checkbox"/> No |
| <input type="checkbox"/> MOP<br><input checked="" type="checkbox"/> MOE<br><br><i>Description:</i><br>Knowledge of military pay should be greater after a L.I.N.K.S. workshop compared with before                             | <input type="checkbox"/> Continuous comparison over time<br><input checked="" type="checkbox"/> Pre/post comparison<br><input type="checkbox"/> Comparison between groups<br><input type="checkbox"/> Comparison with metric<br><br><i>Description:</i><br>Performance on quiz about military pay was compared directly before L.I.N.K.S. workshop and directly after workshop, among workshop participants         | Quiz performance before and after L.I.N.K.S. workshop did not differ significantly    | <i>What is the trend?</i><br><input type="checkbox"/> Better<br><input checked="" type="checkbox"/> Same<br><input type="checkbox"/> Worse | <i>Did this meet your expectations for the program?</i><br><input type="checkbox"/> Met<br><input checked="" type="checkbox"/> Missed<br><input type="checkbox"/> Exceeded | <input checked="" type="checkbox"/> Yes<br><input type="checkbox"/> No |
| <input type="checkbox"/> MOP<br><input checked="" type="checkbox"/> MOE<br><br><i>Description:</i><br>Perceived social support from other military families should be greater after a L.I.N.K.S. workshop compared with before | <input type="checkbox"/> Continuous comparison over time<br><input checked="" type="checkbox"/> Pre/post comparison<br><input type="checkbox"/> Comparison between groups<br><input type="checkbox"/> Comparison with metric<br><br><i>Description:</i><br>Scores on a scale of perceived social support were compared directly before L.I.N.K.S. workshop and 3 months after workshop, among workshop participants | Perceived support was significantly greater after (vs. before) workshop               | <i>What is the trend?</i><br><input checked="" type="checkbox"/> Better<br><input type="checkbox"/> Same<br><input type="checkbox"/> Worse | <i>Did this meet your expectations for the program?</i><br><input checked="" type="checkbox"/> Met<br><input type="checkbox"/> Missed<br><input type="checkbox"/> Exceeded | <input type="checkbox"/> Yes<br><input checked="" type="checkbox"/> No |

example, if 100 percent of L.I.N.K.S. staff need to receive training, you would simply compare the proportion that actually received training to this metric (see second example in Worksheet 6.1).

### Limitations and Potential Sources of Misinterpretation

The data collection and analysis processes that we have described in this chapter offer valuable ways of analyzing MOPs and MOEs to assess how well a PSC is working, but there are inherent limitations to the conclusions that can be drawn from them. The evaluation strategies that we have suggested as most pragmatic and have provided guidance for rely on correlations rather than testing the causal effect of PSCs. Thus, as we have cautioned previously when describing the various analyses, it is not possible to draw conclusions about whether a PSC led to certain outcomes. Rather, we can draw conclusions about associations between different factors that are suggestive of potential program effects. Sources of potential misinterpretation may include values that differ only by chance, co-occurring sources of change, pre-existing differences between respondents, participants who do not respond to a survey, and biased survey responses.

In Table 6.1, we present these sample sources of misinterpretation that may come up when analyzing MOPs and MOEs. We discuss suggestions for recognizing potential biases and drawing accurate conclusions that acknowledge the limitations of what the analyses can speak to.

### A Caution About Overinterpreting *p*-Values

Even when the potential sources for misinterpretation noted in the table have been accounted for, and you achieve a significant result for your PSC, one additional caution should be

**Table 6.1**  
**Sample Sources of Misinterpretation**

| Source of Misinterpretation              | Example                                                                                        | Explanation and Suggestions for Accurate Interpretation                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|------------------------------------------|------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Values that may differ only by chance    | The average satisfaction with FMEAP was 3.6 out of 5 in March and 3.8 in April ( $p = 0.49$ ). | Conducting a statistical test, as described in this chapter and in the RAND Tool, can help to assess whether a difference in values is statistically meaningful. A $p$ -value is implemented to ensure that you do not interpret findings that may be due to chance rather than a true significant change. Throughout this chapter, we discussed using a $p$ -value of 0.05. Here, with a $p$ -value of 0.49, it is clear that we should not place too much stock in the difference between 3.6 and 3.8, although this increase could be a small part of a longer-term gradual increase. Looking at trends over a longer period of time can also help to assess whether a difference in values was merely a fluctuation or part of a broader trend.                                                                                                                                                                                                                                                                                                                                                                                                            |
| Failing to interpret a meaningful effect | The average satisfaction with FMEAP was 2.1 out of 5 in March and 4.8 in April ( $p = 0.10$ ). | As described above, it is important to not overinterpret findings that may be due to chance. However, we also caution against placing too much emphasis on the exact threshold of $p < 0.05$ as an indicator of whether a difference in values is meaningful. There are a number of factors that influence your ability to detect a significant effect, including the strength of the effect and the sample size. The stronger the effect, the smaller the sample needed to detect the effect. Sometimes you may be constrained in the number of people you can include in your evaluation. In this case, if your PSC has a moderate effect, the results of your statistical testing may be nonsignificant, as in this example. However, if you look at the raw values, you see that there has been substantial change in satisfaction levels. Therefore, it may be that the sample was too small to detect the change. In this case, you also have to consider the context for these results—though it may not have reached the threshold of statistical significance, an improvement in satisfaction of this magnitude is important information for the PSC. |

Table 6.1—Continued

| Source of Misinterpretation                                          | Example                                                                                                                                                                   | Explanation and Suggestions for Accurate Interpretation                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|----------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Co-occurring sources of change                                       | The average score on a knowledge quiz regarding understanding of basic military knowledge is greater a week after a L.I.N.K.S. session than before.                       | Depending on when the postsession follow-up survey takes place, individuals who participated in the session may have been exposed to other workshops and social interactions that may also have been sources of basic military knowledge. It is difficult to isolate a PSC as the source of change when other unknown influences are also at play over time. In this case, rather than being able to say that L.I.N.K.S. led to greater knowledge, we could say that individuals showed greater knowledge a week later and that the L.I.N.K.S. session may have been one source of change.                                                                                                                                                                                                     |
| Pre-existing differences between involved and uninvolved respondents | Individuals who use Outdoor Recreation Equipment Checkout and Equipment Rental have greater community connectedness than individuals who do not.                          | Individuals highly involved in a PSC may be individuals who have always been particularly connected to their community, compared with individuals not involved in a PSC. Being active or social may have made them more likely to use the PSC in the first place. Thus, here we could not say that using Outdoor Recreation Equipment Rental leads to greater community connectedness. Instead, it would be more appropriate to say that individuals who use it tend to be higher in community connectedness.                                                                                                                                                                                                                                                                                  |
| Participants who drop out or do not respond                          | The average score on a knowledge quiz is greater after a L.I.N.K.S. session than before, but only half of the participants responded to the postsession follow-up survey. | Individuals who respond to surveys may be different from individuals who do not. In this example, it is possible that average knowledge improved, but it is also possible that participants who did not learn much were not inclined to respond to the follow-up survey, so their lower knowledge did not pull the postsession average downward. Doing a paired t-test among individuals who responded to both surveys is one way to address this issue, but it should be noted that doing so also does not represent the program's average effect across all participants.                                                                                                                                                                                                                    |
| Biased survey answers from respondents                               | Individuals who use FMEAP career coaching indicate on average on a survey that they feel highly prepared for job interviews.                                              | Individuals may feel that they <i>should</i> be highly prepared for job interviews after using career coaching or that the service offering this resource expects them to be more prepared afterward. For sensitive topics, such as relationship skills (a suggested MOE for L.I.N.K.S.), individuals may also feel uncomfortable giving a response that could be judged by the PSC as negative, particularly if they worry that their responses can be linked to their identity. Whether intentional or more subconscious on the part of the respondent, answers given on a survey are not always accurate. Describing differences in values obtained through a survey as differences in “self-reported” outcomes is one way of being clear about the potential biases that could be at work. |

observed: *p*-values indicate statistical significance, but they do not indicate practical significance. Statistical tests in general, and tests using dichotomous variables in particular, are sensitive to the number of observations used in the analysis. So, if you have a large number of participants providing data for the test, then a small difference in the outcome being assessed may be statistically significant, but the difference may have no practical meaning for your PSC. For example, if you had 10,000 participants who completed a survey at two time points, and 50 percent were satisfied or very satisfied with your program at time 1, while 53 percent were satisfied or very satisfied with your program at time 2, that difference would be significant at  $p < 0.05$ . However, it is unclear whether a 3-percent increase in satisfaction has any practical significance for your program. It is important to think carefully and critically about the meaning

of any statistically significant differences you observe in your data, taking into account the size of the difference (e.g., 3 percent is not a large difference), as well as historical trends (e.g., has satisfaction been steadily increasing over time, so that a 3-percent increase fits the trend?) and general variation in the data (e.g., does the program have a well-established history of always having 50-percent satisfaction, so that any change would be meaningful, or do satisfaction ratings vary greatly, suggesting that a 3-percent change is just random variation?).

### ***Considering Generalizability***

Finally, it is important to keep in mind that a significant result in one setting or among one group of people would not necessarily be found in a different setting or a broader population. For example, finding an increase in satisfaction with a PSC on one installation may not mean that there was also an increase on other installations. Thus, caution should be taken to not assume that a finding will generalize to the full population of Marines and families, based on findings that were not specifically designed to represent that full population. It can be helpful to note the characteristics of a given setting or group of participants for a finding, such as that the installation is smaller or the participants are younger. While it is impossible to take all characteristics into account, this step can give a sense of the other contexts to which the findings could potentially generalize.

## **Summary**

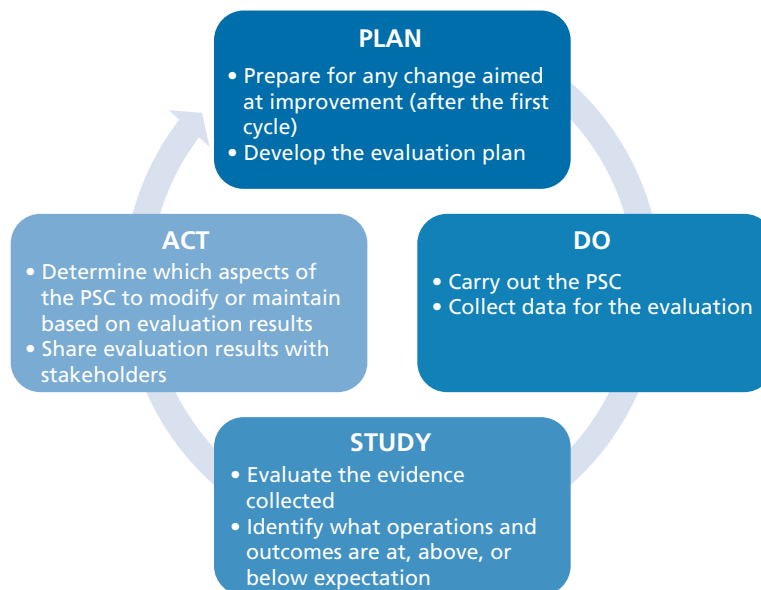
This chapter discussed procedures for data collection and preparing your data for analysis. It then introduced you to the RAND Tool and how this tool can be used to analyze data from your evaluation. We reviewed the use of dichotomous versus continuous data and introduced procedures for analyzing data over time, comparing pre-post data, and comparing PSC participants to members of a comparison group. We also described caveats to keep in mind as you interpret the results of your analyses. The next and final chapter of this user's guide demonstrates how program evaluations can be a part of continuous quality improvement efforts.

## Continuous Quality Improvement

In Chapter One of this user's guide, we explained that program evaluations can be conducted as one-time occurrences but that the more common purpose of program evaluation is as part of continuous improvement efforts. Continuous Quality Improvement (CQI) is a cycle of planned, formal, systematic, and perpetual action for improving your PSC to meet desired outcomes (Berwick, 1989; Deming, 1994; Hunter et al., 2015). Planned refers to the process being a strategic, organized, and intentional investigation into PSC operations rather than an investigation that either occurs as an afterthought (e.g., resources were available at the end of an operations cycle) or relies only on data collected for nonevaluation purposes (e.g., secondary data as discussed in Chapter Five). As demonstrated throughout this User's Guide, evaluation work requires careful consideration of what is being examined, the types of data necessary to measure both performance and effectiveness, and whether the evaluation results suggest the program is meeting its goals or has room for improvement.

But what happens once program staff members know how their PSC's current implementation and effectiveness are associated with desired levels of performance? We suggest following the *Plan-Do-Study-Act (PDSA) Cycle* (shown in Figure 7.1), a modern evolution of CQI, to

**Figure 7.1**  
**Plan-Do-Study-Act Cycle**



RAND TL264-7.1

frame our discussion of what comes next (Deming, 1994; Moen and Norman, 2010). Program evaluation informs all aspects of this cycle. What we have presented so far in this user's guide comprises part of the "Plan" cycle and both the "Do" and "Study" phases of the cycle: Program managers have identified appropriate measures for assessing performance and effectiveness, operated their respective PSC, collected data necessary for the evaluation, and analyzed the data to understand current performance quality. What this chapter focuses on are the "Act" and future "Plan" phases to complete the cycle. By focusing on these two phases of the cycle, we demonstrate how program staff integrates continuous improvement activities into everyday PSC operations.

## Plan-Do-Study-Act Cycle

The PDSA cycle focuses on four distinct phases of action. The plan phase is the preparation for putting planned changes into place (e.g., modifying the L.I.N.K.S. curriculum) and developing a plan for assessing the new initiative or how PSC modifications result in changes to performance. This phase also includes identifying the goals or purposes of your current PDSA cycle, identifying metrics and methods of collecting the data necessary to study changes in the PSC, and defining benchmarks for how success from those changes would look. The do phase is the enacting of the PSC and collecting of the evaluation data. Following this, the study phase is where performance measures and metrics are monitored and analyzed to determine areas of success, areas with room for growth, and areas of struggle that need to be addressed. Finally, the act phase completes the improvement cycle by identifying appropriate actions or changes that will address performance issues in future operation of the program. Once the cycle has concluded, your team returns to the planning phase for the next improvement cycle. We begin our presentation of this cycle by starting at the end: the act phase.

### ACT on Your Evaluation Findings

With evaluation results in hand and an understanding of current areas of practice that are underserved or performing at or above expectation, your team can determine where to focus efforts to improve the delivery of your PSC, with the ultimate intent of improving effectiveness.

Figure 7.2 briefly presents how your team can use the results of both the implementation and outcome evaluations to collectively determine appropriate improvement strategies on which to act.

Continuing with one of the examples used throughout this user's guide, you have now conducted a full evaluation of L.I.N.K.S. and, hypothetically, learned that Marine spouses are not demonstrating the desired level of knowledge related to military pay and promotion practices. Your team assessed this using a knowledge tests developed based on the program's content (see Chapter Five of this user's guide for a refresher on this type of MOE). When examining the implementation fidelity of this program, your team identified two issues: (1) only 75 percent of L.I.N.K.S. trainers had received the required annual training that provides updates on program curricula, and (2) the program curricula had not been updated in the last two years (during which a change to USMC pay processes occurred). Using Figure 7.2, these results would land in the bottom right quadrant: neither implementation fidelity (measured by MOPs) nor desired outcomes (measured by MOEs) were demonstrated to meet desired levels.

**Figure 7.2**  
**Identifying Approaches for Improvement**  
*Adapted from Acosta et al., 2013*

|                                                                                 |     | PSC activities resulted in intended outcomes or significant, positive changes in desired outcomes.                                                                                                                                                                                                                                                                                                                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|---------------------------------------------------------------------------------|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                                 |     | Yes                                                                                                                                                                                                                                                                                                                                                                                                                           | No                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| PSC activities were implemented with fidelity and reached the target population | Yes | <p>The PSC seems to be working as designed; continue implementation and evaluation.</p> <ul style="list-style-type: none"> <li>Continue to monitor process and outcomes</li> </ul>                                                                                                                                                                                                                                            | <p><i>Changes may be needed</i> because your PSC did not achieve the intended results. Lack of significant results could be related to the following barriers:</p> <ul style="list-style-type: none"> <li>A misalignment between PSC activities and intended outcomes.</li> <li>A mismatch between the PSC and its participants (e.g., participants started high on outcome measures, leaving limited room for improvement).</li> <li>A mismatch between evaluation design or measures and PSC outcomes.</li> <li>A small sample size (resulting in lack of power to detect significant effect).</li> </ul> <p>Address these potential barriers before implementing the PSC or prior to the next CQI cycle. Addressing these barriers may require you to change your PSC activities and your evaluation design/measures.</p> |
|                                                                                 | No  | <p><i>Changes may be needed</i> to address potential barriers to fidelity and reach of desired participants; you might want to reevaluate the PSC to determine whether results improve.</p> <ul style="list-style-type: none"> <li>Focus on strengthening areas that were not implemented with fidelity.</li> <li>Improve recruiting or retention strategies to ensure that the PSC reaches its target population.</li> </ul> | <p><i>Changes may be needed</i> because the PSC does not seem to be working as designed. Changes to PSC activities and evaluation design may be required. Assess whether:</p> <ul style="list-style-type: none"> <li>There is a mismatch between the PSC and the target population.</li> <li>There are adequate resources to deliver the PSC (e.g., Do facilitators have enough training? Do participants have incentive to attend or utilize the PSC?).</li> </ul>                                                                                                                                                                                                                                                                                                                                                          |

RAND TL264-7.2

According to the program's logic model (see Chapter Three), there is an expectation that using trained staff to provide a strong curriculum will produce the desired outcomes. The program may be able to reach the desired outcomes without making changes to the program's operation. Thus, when reviewing evaluation results, the program's staff identify the following actions to be taken to improve upon the current level of performance and effectiveness: (1) increased staff training, and (2) updating program curricula to align with current USMC policy.

There will be instances in which program evaluations identify many practices or performance areas that need modification or revamping. In making decisions about how to act or

where to focus change efforts when faced with multiple issues, Hunter and colleagues (2015) recommend considering the following:

1. **Need:** What are the most pressing program issues that need to be addressed?
2. **Capacity:** To what extent does your team have the resources, expertise, and authority to make any necessary modifications to PSC inputs or activities?
3. **Impact:** What changes will lead to the biggest effect on the quality of the program?
4. **Evidence-base:** What actions will help align your PSC with the recommendations in this guide or with prior research on successful PSCs that serve similar communities or have similar goals? Few PSCs in the MF portfolio are found in scholarly literature, but when an evidence-base does exist (for example, see Lerman et al., 2015), the information available can inform the CQI cycle.
5. **Scope:** How much change can your PSC handle at one time? Can you jump right to major shifts in activities without first addressing minor resource issues? Can making small changes now increase the feasibility of larger changes at a later time?

PSCs that operate across multiple installations may respond to program evaluation results in different ways. Local resources and contextual factors likely play a role in determining what actions are feasible for staff to implement (see Chapter Three of this user's guide for information on the role local context plays in PSC operation). That said, consulting with program managers across installations may provide some guidance on best practices or improved resource utilization, reducing the burden on staff for identifying and implementing necessary changes.

Staff should consider creating a tracking system for each aspect of a PSC for which changes, modifications, or improvements are needed. Document the area of deficiency, including potential influencing factors; whether this is a priority area; the action to be taken; the time line or deadline for when changes are expected to be ready or implemented; and who is responsible for leading changes in each area (see an example in Table 7.1). This tracking sheet can build directly from Worksheet 6.2 in this User's Guide. Any MOP or MOE identified as needing action can inform the Performance Area documented in the first column of Worksheet 7.1. Remember to consider contextual factors that may have influenced your PSC's operation and effectiveness when choosing whether to take action and what that action will be.

Revisit this tracking system regularly to ensure the appropriate actions are being taken by the expected deadlines so that changes are instituted before the next evaluation begins.

**Table 7.1**  
**Tracking Areas Identified for Program Improvement**

| Performance Area              | Priority Level | Action(s) to Be Taken                                                                        | Rationale (if No Action) | Expected Completion | Person or Group Responsible |
|-------------------------------|----------------|----------------------------------------------------------------------------------------------|--------------------------|---------------------|-----------------------------|
| L.I.N.K.S. Curriculum Updates | High           | Curriculum edited to include updated policies and ensure complete and accessible information |                          | October 1           | Program Manager             |
| L.I.N.K.S. Trainer Training   | High           | Develop training session for new curriculum materials                                        |                          | November 1          | Program Manager             |
|                               |                | Ensure all LINKS trainers receive training on updated curriculum                             |                          | January 1           |                             |



upcoming PDSA cycle, expected completion dates, and the individuals or group of individuals tasked with the specific actions for each performance area that requires action.

As a final note in the act phase, sharing the results of the evaluation with leadership, stakeholders, and community members can and should occur during the act phase of this cycle. Deciding the specific timing will be dependent on the timing of the evaluation cycle, the results of your evaluation, and the expectations of organization leaders. On one hand, PSC staff may see the need for rapid communication of results to ensure consistent or additional funding in the next operations cycle. On the other hand, PSC staff may want to prepare program improvement plans to accompany evaluation results before sharing what was learned (i.e., waiting until the end of the act phase). Regardless, dissemination of results is important for a number of reasons: this may create strong publicity that reaches your PSC's target audience and assists in future recruitment, may help garner additional support in the shape of funding or staffing for your PSC, and also may help shape policy and practice that informs future PSC efforts (Centers for Disease Control and Prevention, 2009). Finally, sharing results may also provide your staff with an outlet for seeking feedback from your target audience or past participants to guide program improvement efforts.

### **Plan for Changes and Evaluation**

The act phase of the PDSA cycle informs the next plan phase. During this phase of the next cycle, staff members take on two tasks. The first is to prepare or institute the changes staff identified during the prior act phase. Using the example from earlier, L.I.N.K.S. program staff would need to update the curricula to cover current USMC policies and practices. Then, all L.I.N.K.S. trainers would need to be trained on the updated curricula. Second, and this is where program evaluation as part of CQI cycles departs from stand-alone, one-time occurrences, the plan phase is used to prepare for a follow-up evaluation. More specifically, the team should prepare to study whether the chosen modifications create the desired changes in your PSC's performance (e.g., to what extent did improving staff training and updating program curricula improve Marine spouse understanding of pay and promotion procedures?). This means returning to the start of Chapter Four of this user's guide and cycling through it again: designing the appropriate evaluation, identifying measures of performance and effectiveness, collecting and analyzing data, and determining how to respond to the evaluation results to prepare for yet another PDSA cycle.

The plan phase is still active for PSCs in the top-right quadrant of Figure 7.2. Even the highest-quality programs, meeting all outcomes and benchmarks, can benefit from improvement cycles. Traditional CQI encourages the constant questioning of established practice or perspectives such that new or innovative approaches create opportunities for growth (Dean and Mank, 1997). For high-quality PSCs, the act phase is for the exploration of new approaches or considering ways to expand the current PSC's activities. The plan phase is where the staff takes the necessary steps to implement those approaches.

The recently concluded PDSA cycle may also highlight components of a PSC (i.e., resources, activities, or outputs) where little information is currently available on performance. Maybe the MOPs did not capture what your team expected or were not fine-grained (i.e., specific) enough to give you the level of detail necessary to identify how well the program is being implemented. The act phase can be used to assess issues with prior evaluations and identify "improving the evaluation" as the desired change to operations in future continuous improvement cycles. Processing through the plan phase of the PDSA cycle allows your team to prepare

for new collection measures or develop new assessment plans to address knowledge gaps and improve upon past program evaluation efforts.

### **Keep the Cycle Going**

CQI is a practice of monitoring current performance, enacting corrective action on or modifications to specific aspects of a program, and assessing subsequent changes in program performance and outcomes. Staff can use a single PDSA cycle to implement large changes or focus on small modifications to only one aspect of a PSC. However, true continuous quality improvement incorporates planned, formal, and systematic evaluations of the day-to-day operations of the program for the cycle of improvement to be in perpetual action (Ryan et al., 2014).

Each PDSA cycle may focus on different aspects of a PSC rather than examining all components of the program's implementation and effectiveness in each cycle. Due to resource limitations and time demands, evaluations will likely never be able to answer all questions about a program's operation and effectiveness. Your team can prioritize specific aspects of the PSC to focus on based on the theory of change identified in the logic model, where a program is in its own development, or the intended use of the evaluation (National Center for Chronic Disease Prevention and Health Promotion, 2011). For example, if a change is being made to only one activity within the PSC, the team may focus only on that activity and its associated outputs and outcomes. If a PSC is new, the evaluation team may focus more on process, especially over intermediate and long-term outcomes. Finally, a policy or regulation change may prioritize particular components of a PSC (e.g., L.I.N.K.S. for dependent children) over others so that the necessary evidence is available when responding to the new policy. Advance planning of evaluations allows staff to present a rationale for the limited scope while maintaining support from key stakeholders and decisionmakers. When a team does elect to focus only on one particular aspect of a PSC, it should consider collecting all data measures as though a full evaluation were being conducted; these data points may be critical in future PDSA cycles when the focus is shifted to other areas of the operation and comparison over time measures are needed.

A final note about evaluation cycles: The evaluation team should avoid getting stuck in one particular phase without advancing. Good evaluations collect data, analyze data, and report results quickly to implement changes quickly; they allow participants and stakeholders to be informed of findings in a timely manner to emphasize why participating or contributing was important; and they demonstrate that the results are being used to make change so that continued buy-in from installations and personnel remains strong or improves. That said, each phase of the cycle may require different time allotments. A well-designed evaluation cycle should account for the amount of time between a program's operation completing and when identified changes need to be ready so they can be incorporated in future operating cycles. Knowing these deadlines in advance may help keep the cycle moving forward.

**Info Box 7.1****Program Evaluation and Strategic Portfolio Management (SPM)**

As demonstrated throughout this toolkit, program evaluation is a process for systematically collecting and analyzing performance and outcome data on a single PSC. The process is used to identify whether a PSC is meeting intended goals and where performance can improve with modification or additions to current PSC inputs and activities. Program evaluation has a narrow focus on performance and effectiveness, relative to the goals served by each MF branch and the entire MF Division.

Each MF branch manages a portfolio of PSCs meant to work collectively to meet the broad purposes of that branch, such as increasing retention or improving resiliency of active-duty Marines. Each branch must balance the types of PSCs, what each PSC is expected to do for the branch, and how resources get allocated to each individual PSC. Each branch functions at the installation or force level, moving beyond the scope of work served by each individual PSC at an installation. Strategic management outlines a process that guides the portfolio design, provides ongoing management of the portfolio, and helps to determine resource allocation across a set of PSCs (Poister and Streib, 1999). Portfolio evaluation is a critical part of this management process and can be used to assess whether the collective set of PSCs in the portfolio is being well managed, whether the portfolio contains PSCs that are meaningfully connected, and whether the portfolio is agile, addressing new needs of the force as they arise.

Portfolio evaluation can assist branches in determining whether the programs, working collectively, are meeting long-term objectives (e.g., improving retention). Portfolio evaluation is not as simple as aggregating data and measures of performance and effectiveness from each individual PSC. In fact, relying on PSC evaluation data alone can be burdensome, if not impossible, for branches seeking to integrate and form a comprehensive portfolio evaluation (Reed and Jordan, 2007).

Branch and MF managers need to consider “agency-level” measures that can identify the extent to which the portfolio of PSCs is interacting to produce the desired changes to expected areas of impact (Reed and Jordan, 2007). No single PSC is solely responsible for the changes to these impact areas. As an example, if we look at one possible area where the MF PSCs are expected to make an impact—Marine and family resilience—there are at least 32 individual PSCs within the MF portfolio that are intended to help address this area. They represent each of the five branches used to guide the development of this toolkit and include PSCs such as the DSTRESS hotline, Child and Youth Programs, L.I.N.K.S., FMEAP, and the Single Marine Program.

In the same way that Chapter Five of this toolkit discusses the development of MOEs for short- and intermediate-term outcomes, branch managers and the MF Division should work collaboratively to determine the appropriate way to assess each area of impact. This includes whether measures will be assessed as changes over time, differences between two groups (e.g., Marines who do and do not engage in MF PSCs), or as a change within each individual Marine and military dependent who participates in MF PSCs. Further consideration should be given to whether these effects are measured for single installations or a combination of some/all installations, particularly because Marines may interact with PSCs at multiple locations that intend to address impacts of interest during the time frame of the evaluation. Importantly, if individual PSC staff will be involved in the collection of data for the measures of effectiveness, it is critical for the staff overseeing the portfolio impact evaluation to standardize what information is required to be collected, as well as how and when PSC staff should collect it. Portfolio evaluation depends on standard measurement at the individual level so that those portfolio impacts can be evaluated at the aggregate level.

Providing a guide for strategic portfolio management is outside the scope of this toolkit, which is focused on program evaluation. However, when MF or individual branches determine that desired impacts are being achieved, staff can consider how to further maximize resource allocation or whether there are new needs within the Marine Corps that PSCs should now address. If an evaluation demonstrates that the collective set of PSCs are not driving changes in desired areas of impact, branch and MF staff need to determine where there are operational gaps that can be addressed. Maybe a PSC needs to repurpose some of its programmatic efforts, or maybe an additional PSC is needed to help the portfolio reach its goals. Just as program evaluation is a component of a larger improvement and operations cycle, portfolio evaluation should be a regular aspect of managing a portfolio of PSCs that serve our military community.

**Conclusion**

We hope you found the information provided helpful for planning, implementing, and using the results of your evaluation for program improvement. If, at any point, you skipped any of the steps or worksheets provided in this user’s guide, we recommend revisiting those sections. Each chapter builds on the information and worksheets developed in prior chapters.

- Chapter One offered an overview of program evaluation and the purpose and layout of the user’s guide.

- Chapter Two provided a framework for understanding whether your team works with a program, service, or capability.
- Chapter Three helped you identify core components of your PSC and organize these components into a logic model. This logic model then guides your evaluation planning.
- Chapter Four described the design, focus, and types of process and outcome evaluations that your team could consider using in an evaluation of your PSC.
- Chapter Five presented measures of performance and measures of effectiveness that link to the short- and intermediate-term outcomes identified in the development of your logic model.
- Chapter Six explained potential tools or processes to collect, organize, and analyze your evaluation data and showed how to interpret evaluation results.
- Chapter Seven demonstrated how program evaluation results guide continuous program improvement efforts.

By completing each chapter, your evaluation team can be certain that you have taken advantage of the complete package of information, insight, and guidance provided in this user's guide. As suggested earlier in this chapter, this user's guide is only the start of your program evaluation work. Continuous improvement operates in a cycle of planning for evaluation, operating your PSC and collecting evaluation data, studying the data collected, and acting on those results to improve PSC operation and PSC effectiveness. Each complete CQI cycle informs the next cycle's planning phase. If your team plans for changes to a PSC's inputs, activities, or outputs, your team will need to return to Chapter Three, modify the PSC's logic model, and complete the remainder of the user's guide for your next evaluation cycle. If your team identifies process improvements to current PSC activities, then your team can return to Chapter Four of this user's guide to plan the next evaluation, then progress through Chapters Five through Seven to complete the next evaluation cycle. The user's guide was developed so that your team can rely on it to guide each new cycle of program improvement and to continually enhance your team's evaluation efforts. To that end, we hope you have found and continue to find the user's guide, and associated Workbooks, to be a helpful resource.



## References

---

- Acosta, J. D., R. N. Ramchand, A. Becker, A. Felton, and A. Kofner, *RAND Suicide Prevention Program Evaluation Toolkit*, Santa Monica, Calif.: RAND Corporation, TL-111-OSD, 2013. As of January 11, 2021: <https://www.rand.org/pubs/tools/TL111.html>
- Bandura, A., “Guide for Constructing Self-Efficacy Scales,” *Self-Efficacy Beliefs of Adolescents*, Vol. 5, 2006, pp. 307–337.
- Berwick, D. M., “Continuous Improvement as an Ideal in Health Care,” *New England Journal of Medicine*, Vol. 320, No. 1, 1989, pp. 53–56.
- Centers for Disease Control and Prevention, *Disseminating Program Achievements and Evaluation Findings to Garner Support*, U.S. Department of Health and Human Services, 2009. As of January 11, 2021: <https://www.cdc.gov/healthyyouth/evaluation/pdf/brief9.pdf>
- , *Introduction to Program Evaluation for Public Health Programs: A Self-Study Guide*, U.S. Department of Health and Human Services, 2011.
- Chang, A., and P. Bordia, “A Multidimensional Approach to the Group Cohesion–Group Performance Relationship,” *Small Group Research*, Vol. 32, 2001, pp. 379–405.
- Chinman, M., P. S. Imm, and A. H. Wandersman, *Getting to Outcomes™ 2004: Promoting Accountability Through Methods and Tools for Planning, Implementation, and Evaluation*, Santa Monica, Calif.: RAND Corporation, TR-101-CDC, 2003. As of January 11, 2021: [https://www.rand.org/pubs/technical\\_reports/TR101.html](https://www.rand.org/pubs/technical_reports/TR101.html)
- Code of Federal Regulations, Title 45, Public Welfare, Department of Health and Human Services, Part 46: Protection of Human Subjects, revised January 15, 2009, effective July 14, 2009.
- Cook, T. D., and D. T. Campbell, *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Boston, Mass.: Houghton Mifflin, 1979.
- Daugherty, L., L. Werber, K. Kamarck, L. Harrington, and J. Gazis, *Officer Accession Planning: A Manual for Estimating Air Force Officer Degree Requirements*, Santa Monica, Calif.: RAND Corporation, TL-196-AF, 2016. As of January 11, 2021: <https://www.rand.org/pubs/tools/TL196.html>
- Dean, J. E. A., and D. M. Mank, “Continuous Improvement and Quality of Life: Lessons from Organizational Management,” in R. L. Schalock and G. N. Siperstein, eds., *Quality of Life Volume II: Application to Persons with Disabilities*, Washington, D.C.: American Association on Mental Retardation, 1997.
- Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury, *Program Evaluation Guide*, Arlington, Va.: Defense Health Agency, 2012.
- Department of Defense Instruction 3216.02, *Protection of Human Subjects and Adherence to Ethical Standards in DoD-Supported Research*, November 8, 2011.
- Department of Defense Instruction 8550.01, *DoD Internet Services and Internet-Based Capabilities*, September 11, 2012.
- Deming, W. E., *The New Economics for Industry, Government, Education*, Cambridge, Mass.: Massachusetts Institute of Technology, Center for Advanced Educational Services, 1994.

Fitzpatrick, J. L., J. R. Sanders, and B. R. Worthen, *Program Evaluation: Alternative Approaches and Practical Guidelines*, 4th ed., Boston, Mass.: Pearson, 2011.

Gottman, J. M., *What Predicts Divorce? The Relationship Between Marital Processes and Marital Outcomes*, Hillsdale, N.J.: Lawrence Erlbaum Associates, 1994.

———, *The Relationship Clinic: A Scientifically Based Marital Therapy*, New York: W. W. Norton & Company, 1999.

Holliday, S., S. L. Wrabel, T. Trail, M. Tankard, D. Schultz, L. Werber, and S. O. Meadows, *A Guide to Evaluating Marine and Family Division Programs, Services, and Capabilities: An Introduction to Program Evaluation*, Santa Monica, Calif.: RAND Corporation, TL-264-USMC/1, 2021. As of January 11, 2021: <http://www.rand.org/pubs/tools/TL264z1.html>

Hourcade, J., P. Parette, and H. Anderson, "Accountability in Collaboration: A Framework for Evaluation," *Education and Training in Developmental Disabilities*, Vol. 38, No. 4, 2003, pp. 398–404.

Hox, J. J., and H. R. Boeije, "Data Collection, Primary Versus Secondary," in K. K. Leonard, ed., *Encyclopedia of Social Measurement*, Amsterdam, Netherlands: Elsevier, 2005, pp. 593–599.

Hunter, S. B., P. A. Ebener, M. Chinman, A. J. Ober, and Christina Y. Huang, *Promoting Success: A Getting to Outcomes® Guide to Implementing Continuous Quality Improvement for Community Service Organizations*, Santa Monica, Calif.: RAND Corporation, TL-179-NIDA, 2015. As of January 11, 2021: <http://www.rand.org/pubs/tools/TL179.html>

Lalayants, M., "Differential Program Evaluation Model in Child Protection," *Child Welfare*, Vol. 91, No. 4, 2012, pp. 9–40.

Lawton, B., P. R. Brandon, L. Cicchinelli, and W. Kekahio, *Logic Models: A Tool for Designing and Monitoring Program Evaluations* (REL 2014–007), Washington, D.C.: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Pacific, 2014. As of January 11, 2021: [https://ies.ed.gov/ncee/edlabs/regions/pacific/pdf/REL\\_2014007.pdf](https://ies.ed.gov/ncee/edlabs/regions/pacific/pdf/REL_2014007.pdf)

LeBlanc, P., and C. Lacey, *Peace Works: A Program Evaluation Model*, paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, La., 2002.

Lerman, R. A., D. Hanson, M. Tanamor, L. Blatt, and C. Dobson, *The United Services Military Apprenticeship Program (USMAP): Implementation Study and Feasibility of an Impact Study*, Washington, D.C.: Urban Institute, 2015. As of January 11, 2021: [https://www.dol.gov/asp/evaluation/completed-studies/The\\_United\\_Services\\_Military\\_Apprenticeship\\_Program\\_\(USMAP\).pdf](https://www.dol.gov/asp/evaluation/completed-studies/The_United_Services_Military_Apprenticeship_Program_(USMAP).pdf)

Lesesne, C. A., K. M. Lewis, C. Moore, D. Fisher, D. Green, and A. Wandersman, *Promoting Science-Based Approaches to Teen Pregnancy Prevention Using Getting to Outcomes: Draft 2011*, unpublished manual, 2011.

Marine Corps Order 3900.18, *Human Research Protection Program (HRPP)*, January 21, 2011.

Marine Corps Order 5300.18, *Marine Corps Survey Program*, May 21, 2015.

Martin, L. T., C. Farris, D. M. Adamson, and R. M. Weinick, *A Systematic Process to Facilitate Evidence-Informed Decisionmaking Regarding Program Expansion: The RAND Toolkit, Volume 3*, Santa Monica, Calif.: RAND Corporation, RR-487/3-OSD, 2014. As of January 11, 2021: [https://www.rand.org/pubs/research\\_reports/RR487z3.html](https://www.rand.org/pubs/research_reports/RR487z3.html)

McClure, P., and W. Broughton, "Measuring the Cohesion of Military Families," *Armed Forces and Society*, Vol. 26, 2000, pp. 473–487.

McNeil, R. C., "A Program Evaluation Model: Using Bloom's Taxonomy to Identify Outcome Indicators in Outcomes-Based Program Evaluations," *Journal of Adult Education*, Vol. 40, No. 2, 2011, pp. 24–29.

Moen, R. D., and C. L. Norman, "Circling Back: Clearing up Myths About the Deming Cycle and Seeing How It Keeps Evolving," *Quality Progress*, Vol. 43, 2010, pp. 22–28.

- National Center for Chronic Disease Prevention and Health Promotion, *Developing an Effective Evaluation Plan*, Atlanta, Ga.: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, Division of Nutrition, Physical Activity, and Obesity, 2011. As of January 11, 2021:  
<https://www.cdc.gov/obesity/downloads/cdc-evaluation-workbook-508.pdf>
- National Center for Education Evaluation and Regional Assistance, *Identifying and Implementing Educational Practices Supported by Rigorous Evaluation: A User Friendly Guide*, Washington, D.C.: Institute of Education Sciences, 2003.
- Patton, Q., *Qualitative Research and Evaluation Methods: Integrating Theory and Practice*, 4th ed., Thousand Oaks, Calif.: Sage Publications, 2015.
- Peterson, N. A., P. W. Speer, and D. W. McMillan, "Validation of a Brief Sense of Community Scale: Confirmation of the Principal Theory of Sense of Community," *Journal of Community Psychology*, Vol. 36, 2008, pp. 61–73.
- Poister, T. H., and G. D. Streib, "Strategic Management in the Public Sector: Concepts, Models, and Processes," *Public Productivity & Management Review*, Vol. 22, No. 3, 1999, pp. 308–325.
- Reed, J. H., and G. Jordan, "Using Systems Theory and Logic Models to Define Integrated Outcomes and Performance Measures in Multi-Program Settings," *Research Evaluation*, Vol. 16, No. 3, 2007, pp. 169–181.
- Rossi, P. H., M. W. Lipsey, and H. E. Freeman, *Evaluation: A Systematic Approach*, 7th ed., Thousand Oaks, Calif.: Sage Publications, 2004.
- Ryan, G., C. M. Farmer, D. A. Adamson, and R. M. Weinick, *A Program Manager's Guide for Program Improvement in Ongoing Psychological Health and Traumatic Brain Injury Programs: The RAND Toolkit, Volume 4*, Santa Monica, Calif.: RAND Corporation, RR-487/4-OSD, 2014. As of January 11, 2021:  
[https://www.rand.org/pubs/research\\_reports/RR487z4.html](https://www.rand.org/pubs/research_reports/RR487z4.html)
- Secretary of the Navy Manual 5210.1 *Department of the Navy Records Management Program Records Management Manual*, September 15, 2017, Chapter 1.
- Secretary of Navy Instruction 5211.5E, *Department of the Navy Privacy Act Program*, December 28, 2005
- Secretary of Navy Instruction 3900.39D, *Human Research Protection Program*, November 6, 2006.
- Shadish, W. R., T. D. Cook, and D. T. Campbell, eds., *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, 2nd ed., Boston, Mass.: Houghton Mifflin, 2002.
- Shadish, W. R., T. D. Cook, and L. C. Leviton, *Foundations of Program Evaluation: Theories of Practice*, Newbury Park, Calif.: Sage Publications, 1991.
- Smith, M. D., and J. D. Hagman, *Year 1 Assessment of the Unit Focused Stability Manning System*, final report, U.S. Army Research Institute for the Behavioral and Social Sciences, 2004. As of January 11, 2021:  
<https://apps.dtic.mil/sti/citations/ADA428049>
- Stith, S. M., and S. L. Hamby, "The Anger Management Scale: Development and Preliminary Psychometric Properties," *Violence and Victims*, Vol. 17, No. 4, 2002, pp. 383–402.
- Teddlie, C., and F. Yu, "Mixed Methods Sampling: A Typology with Examples," *Journal of Mixed Methods Research*, Vol. 1, No. 1, 2007, pp. 77–100.
- Unger, D. G., E. A. Park, P. Antal, P. A. Tressell, K. Rigney, K. DeRasmo, and J. Kassess, "Serving Children with Special Social and Emotional Needs: A Practical Approach to Evaluating Prevention Programs in Schools and Community Settings," *Journal of Educational & Psychological Consultation*, Vol. 11, No. 2, 2000, pp. 273–296.
- U.S. Marine Corps, Marine Corps Human Research Protection Program (HRPP) and Institutional Review Board (IRB), website, undated. As of January 11, 2021:  
<https://www.tecom.marines.mil/Resources/Marine-Corps-Human-Research-Protection-Program/>

Weinick, R. M., E. B. Beckjord, C. M. Farmer, L. T. Martin, E. M. Gillen, J. D. Acosta, M. P. Fisher, J. Garnett, G. C. Gonzalez, T. C. Helmus, L. H. Jaycox, K. A. Reynolds, N. Salcedo, N., and D. M. Scharf, *Programs Addressing Psychological Health and Traumatic Brain Injury Among U.S. Military Service Members and Their Families*, Santa Monica, Calif.: RAND Corporation, TR-950-OSD, 2011. As of January 11, 2021: [https://www.rand.org/pubs/technical\\_reports/TR950.html](https://www.rand.org/pubs/technical_reports/TR950.html)

Westen, D., and R. Rosenthal, "Quantifying Construct Validity: Two Simple Measures," *Journal of Personality and Social Psychology*, Vol. 84, No. 3, 2003, pp. 608–618.

What Works Clearinghouse, *Procedures and Standards Handbook Version 3.0*, Washington, D.C.: Institute of Education Sciences, 2014.

Wholey, J. S., H. P. Hatry, and K. E. Newcomer, eds., *Handbook of Practical Program Evaluation*, 3rd ed., San Francisco, Calif.: Jossey-Bass, 2010.

W. K. Kellogg Foundation, *W.K. Kellogg Foundation Evaluation Handbook*, Battle Creek, Mich., 2004a.

———, *Logic Model Development Guide*, Battle Creek, Mich., 2004b. As of January 11, 2021: <https://www.wkkf.org/resource-directory/resource/2006/02/wk-kellogg-foundation-logic-model-development-guide>

World Bank, *Doing Impact Evaluation Number 6: Data for Impact Evaluation*, Washington, D.C.: Poverty Reduction and Economic Management, Thematic Group on Poverty Analysis, Monitoring and Impact Evaluation, World Bank, 2009.

World Health Organization, *WHO Evaluation Practice Handbook*, Geneva, Switzerland, 2013.

Zimet, G. D., N. W. Dahlem, S. G. Zimet, and G. K. Farley, "The Multidimensional Scale of Perceived Social Support," *Journal of Personality Assessment*, Vol. 52, No. 1, 1988, pp. 30–41.

For Marine and Family Division (MF) programs and services to remain meaningful, credible, and practical, their effectiveness must be able to be demonstrated. Assessing their offerings will help the U.S. Marine Corps (USMC) to surge and contract the portfolio as resources wax and wane. This program evaluation toolkit is designed to help those responsible for implementing MF programs to determine whether those programs are effective—achieving their intended goals—and ultimately to guide efficient use of scarce resources. This toolkit is composed of two interrelated documents: this User’s Guide and a set of Microsoft Excel workbooks that correspond to material presented in the User’s Guide about collecting and analyzing data. This toolkit is based on the program evaluation literature and program evaluation toolkits produced by other RAND researchers. A review of relevant Marine Corps Orders; discussions with headquarters, branch, and program staff in MF; and reviews of publicly available websites provided additional background for the specific context of MF, USMC. The contents of this toolkit should be of particular interest to MF and staff in the five branches within the scope of this study: Behavioral Health, Family Care, Family Readiness, Personal and Professional Development, and Semper Fit. It may also be of interest to those who manage or oversee similar programs that provide support to service members and their families.



NATIONAL DEFENSE RESEARCH INSTITUTE

[www.rand.org](http://www.rand.org)

\$49.95

ISBN-10 197740769-2  
ISBN-13 9781977407696

