

EVALUATION

A Systematic Approach
SEVENTH EDITION

PETER H. ROSSI

University of Massachusetts, Amherst

MARK W. LIPSEY

Vanderbilt University, Nashville, TN

HOWARD E. FREEMAN



SAGE Publications

International Educational and Professional Publisher
Thousand Oaks ■ London ■ New Delhi

Copyright © 2004 by Sage Publications, Inc.

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

For information:



Sage Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

Sage Publications Ltd.
6 Bonhill Street
London EC2A 4PU
United Kingdom

Sage Publications India Pvt. Ltd.
B-42, Panchsheel Enclave
Post Box 4109
New Delhi 110 017 India

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Rossi, Peter Henry, 1921-

Evaluation : a systematic approach / by Peter H. Rossi, Mark W.

Lipsey, Howard E. Freeman.— 7th ed.

p. cm.

Includes bibliographical references and index.

ISBN 0-7619-0894-3 (cloth)

1. Evaluation research (Social action programs) I. Lipsey, Mark W.

II. Freeman, Howard E. III. Title.

H62.R666 2004

361.6'1'072—dc22

2003015926

Printed on acid-free paper

03 04 05 06 07 08 09 10 9 8 7 6 5 4 3 2 1

<i>Acquiring Editor:</i>	Lisa Cuevas Shaw
<i>Associate Editor:</i>	Paul Reis
<i>Developmental Editor:</i>	John Bergez
<i>Production Editor:</i>	Claudia A. Hoffman
<i>Copy Editor:</i>	Kate Peterson
<i>Typesetter:</i>	C&M Digitals (P) Ltd.
<i>Indexer:</i>	Molly Hall
<i>Cover Designer:</i>	Michelle Lee Kenny

Analysis of Program Process Monitoring Data

Description of the Program Operation

Comparison Between Sites

Conformity of the Program to Its Design

To be effective in bringing about the desired improvements in social conditions, a program needs more than a good plan of attack. Most important, the program must implement its plan; that is, it must actually carry out its intended functions in the intended way.

Although implementing a program concept may seem straightforward, in practice it is often very difficult. Social programs typically must contend with many adverse influences that can compromise even well-intentioned attempts to conduct program business appropriately. The result can easily be substantial discrepancies between the program as intended and the program as actually implemented.

The implementation of a program is reflected in concrete form in the program processes that it puts in place. An important evaluation function, therefore, is to assess the adequacy of program process: the program activities that actually take place and the services that are actually delivered in routine program operation. This chapter introduces the procedures evaluators use to investigate these issues.

After signing a new bill, President John F. Kennedy is reputed to have said to his aides, "Now that this bill is the law of the land, let's hope we can get our government to carry it out." Both those in high places and those on the front lines are often justified in being skeptical about the chances that a social program will be appropriately implemented. Many steps are required to take a program from concept to full operation, and much effort is needed to keep it true to its original design and purposes. Thus, whether any program is fully carried out as envisioned by its sponsors and managers is always problematic.

Ascertaining how well a program is operating, therefore, is an important and useful form of evaluation, known as *program process evaluation*. (A widely used alternative label is *implementation evaluation*.) It does not represent a single distinct evaluation procedure but, rather, a family of approaches, concepts, and methods. The defining theme of program process evaluation (or simply process evaluation) is a focus on the enacted program itself—its operations, activities, functions, performance, component parts, resources, and so forth. When process evaluation involves an ongoing effort to

measure and record information about the program's operation, we will refer to it as *program process monitoring*.

What Is Program Process Evaluation and Monitoring?

As was suggested in Chapter 2, evaluators often distinguish between process (or implementation) evaluation and outcome (or impact) evaluation. Process evaluation, in Scheirer's (1994) words, "verifies what the program is and whether or not it is delivered as intended to the targeted recipients." It does not, however, attempt to assess the effects of the program on those recipients. Such assessment is the province of impact evaluation, which we consider in later chapters.

Where process evaluation is an ongoing function involving repeated measurements over time, it is referred to as program monitoring. Corresponding to the distinction between process and outcome evaluation, **program process monitoring** is the systematic and continual documentation of key aspects of program performance that assesses whether the program is operating as intended or according to some appropriate standard, whereas **outcome monitoring** is the continual measurement of intended outcomes of the program, usually of the social conditions it is intended to improve. We discuss outcome monitoring in conjunction with impact evaluations later in this book.

Program process evaluation generally involves assessments of program performance in the domains of service utilization and program organization. Assessing service utilization consists of examining the extent to which the intended target population receives the intended services. Assessing program organization requires comparing the plan for what the program should be doing with what is actually done, especially with regard to providing services. Usually, program process evaluation is directed at one or both of two key questions: (1) whether a program is reaching the appropriate target population and (2) whether its service delivery and support functions are consistent with program design specifications or other appropriate standards. Process evaluation may also examine what resources are being or have been expended in the conduct of the program.

More specifically, program process evaluation schemes are designed to answer such evaluation questions as these:

- How many persons are receiving services?
- Are those receiving services the intended targets?
- Are they receiving the proper amount, type, and quality of services?

- Are there targets who are not receiving services or subgroups within the target population who are underrepresented among those receiving services?
- Are members of the target population aware of the program?
- Are necessary program functions being performed adequately?
- Is staffing sufficient in numbers and competencies for the functions that must be performed?
- Is the program well organized? Do staff work well with each other?
- Does the program coordinate effectively with the other programs and agencies with which it must interact?
- Are resources, facilities, and funding adequate to support important program functions?
- Are resources used effectively and efficiently?
- Is the program in compliance with requirements imposed by its governing board, funding agencies, and higher-level administration?
- Is the program in compliance with applicable professional and legal standards?
- Is performance at some program sites or locales significantly better or poorer than at others?
- Are participants satisfied with their interactions with program personnel and procedures?
- Are participants satisfied with the services they receive?
- Do participants engage in appropriate follow-up behavior after service?

Setting Criteria for Judging Program Process

It is important to recognize the evaluative themes in process evaluation questions such as those listed above. Virtually all involve words such as *appropriate, adequate, sufficient, satisfactory, reasonable, intended*, and other phrasing indicating that an evaluative judgment is required. To answer these questions, therefore, the evaluator or other responsible parties must not only describe the program's performance but also assess whether it is satisfactory. This, in turn, requires that there be some bases for making judgments, that is, some defensible criteria or standards to apply. Where such criteria are not already articulated and endorsed, the evaluator may find that establishing workable criteria is as difficult as determining program performance on the pertinent dimensions.

There are several approaches to the matter of setting criteria for program performance. Moreover, different approaches will likely apply to different dimensions of program performance because the considerations that go into defining, say, what constitutes an appropriate number of clients served are quite different from those pertinent to deciding what constitutes an adequate level of resources. This said, the approach to the criterion issue that has the broadest scope and most general utility in program process evaluation is the application of program theory as described in Chapter 5.

Recall that program theory, as we presented it, is divided into program process theory and program impact theory. Program process theory is formulated to describe the program as intended in a form that virtually constitutes a plan or blueprint for what the program is expected to do and how. As such, it is particularly relevant to program process evaluation. Recall also that program theory builds on needs assessment (whether systematic or informal) and thus connects the program design with the social conditions the program is intended to ameliorate. And, of course, the process through which theory is derived and adopted usually involves input from the major stakeholders and, ultimately, their endorsement. Program theory thus has a certain authority in delineating what a program “should” be doing and, correspondingly, what constitutes adequate performance.

Program process evaluation, therefore, can be built on the scaffolding of program process theory. Process theory identifies the aspects of program performance that are most important to describe and also provides some indication of what level of performance is intended, thereby providing the basis for assessing whether actual performance measures up. Exhibit 5-E in the previous chapter, for instance, illustrates the service utilization component of the program process theory for an aftercare program for released psychiatric patients. This flowchart depicts, step by step, the interactions and experiences patients released from the hospital are supposed to have as a result of program service. A thorough monitoring procedure would systematically document what actually happened at each step. In particular, it would report how many patients were released from the hospital each month, what proportion were visited by a social worker, how many were referred to services and which services, how many actually received those services, and so forth.

If the program processes that are supposed to happen do not happen, then we would judge the program’s performance to be poor. In actuality, of course, the situation is rarely so simple. Most often, critical events will not occur in an all-or-none fashion but will be attained to some higher or lower degree. Thus, some, but not all, of the released patients will receive visits from social workers, some will be referred to services, and so forth. Moreover, there may be important quality dimensions. For instance, it would not represent good program performance if a released patient was

referred to several community services, but these services were inappropriate to the patient's needs. To determine how much must be done, or how well, we need additional criteria that parallel the information the monitoring procedure provides. If the monitoring procedure reports that 63% of the released patients are visited by a social worker within two weeks of release, we cannot evaluate that performance without some standard that tells us what percentage is "good." Is 63% a poor performance, given that we might expect 100% to be desirable, or is it a very impressive performance with a clientele that is difficult to locate and serve?

The most common and widely applicable criteria for such situations are simply **administrative standards** or objectives, that is, stipulated achievement levels set by program administrators or other responsible parties. For example, the director and staff of a job training program may commit to attaining 80% completion rates for the training or to having 60% of the participants permanently employed six months after receiving training. For the aftercare program above, the administrative target might be to have 75% of the patients visited within two weeks of release from the hospital. By this standard, the 63% found with program monitoring is a subpar performance that, nonetheless, is not too far below the mark.

Administrative standards and objectives for program process performance may be set on the basis of past experience, the performance of comparable programs, or simply the professional judgment of program managers or advisers. If they are reasonably justified, they can provide meaningful standards against which to assess observed program performance. In a related vein, some aspects of program performance may fall under applicable legal, ethical, or professional standards. The "standards of care" adopted in medical practice for treating common ailments, for instance, provide a set of criteria against which to assess program performance in health care settings. Similarly, state children's protective services almost always have legal requirements to meet concerning handling cases of possible child abuse or neglect.

In practice, the assessment of particular dimensions of program process performance is often not based on specific, predetermined criteria but represents an after-the-fact judgment call. This is the "I'll know it when I see it" school of thought on what constitutes good program performance. An evaluator who collects process data on, say, the proportion of high-risk adolescents who recall seeing program-sponsored antidrug media messages may find program staff and other key stakeholders resistant to stating what an acceptable proportion would be. If the results come in at 50%, however, a consensus may arise that this is rather good considering the nature of the population, even though some stakeholders might have reported much higher expectations prior to seeing the data. Other findings, such as 40% or 60%, might also be considered rather good. Only extreme findings, say 10%, might strike all stakeholders as distressingly low. In short, without specific before-measurement criteria a wide range of performance might

be regarded as acceptable. Of course, assessment procedures that are too flexible and that lead to a “pass” for all tend to be useless.

Very similar considerations apply to the organizational component of the process theory. A depiction of the organizational plan for the aftercare program was presented in Exhibit 5-F in Chapter 5. Looking back at it will reveal that it, too, identifies dimensions of program performance that can be described and assessed against appropriate standards. Under that plan, for instance, case managers are expected to interview clients and families, assess service needs, make referrals to services, and so forth. A program process evaluation would document and assess what was done under each of those categories.

Common Forms of Program Process Evaluations

Description and assessment of program process are quite common in program evaluation, but the approaches used are varied, as is the terminology they employ. Such assessments may be conducted as a one-shot endeavor or may be continuous so that information is produced regularly over an extended period of time, as in program process monitoring. They may be conducted by “outside” or “inside” evaluators or be set up as management tools with little involvement by professional evaluators. Moreover, their purpose may be to provide feedback for managerial purposes, to demonstrate accountability to sponsors and decisionmakers, to provide a freestanding process evaluation, or to augment an impact evaluation. Amid this variety, we distinguish two principal forms of program process studies, process or implementation evaluation and continuous program monitoring.

Process or Implementation Evaluation

Process or implementation evaluation is typically conducted by evaluation specialists as a separate project that may involve program personnel but is not integrated into their daily routine. When completed and, often, while under way, process evaluation generally provides information about program performance to program managers and other stakeholders, but is not a regular and continuing part of a program’s operation. Exhibit 6-A describes a process evaluation of an integrated services program for children.

As an evaluation approach, process evaluation plays two major roles. First, it can stand alone as an evaluation of a program in circumstances where the only questions at issue are about the integrity of program operations, service delivery, and other such matters. There are several kinds of situations that fit this description. A stand-alone process evaluation might be appropriate for a relatively new program, for instance, to answer questions about how well it has established its intended operations and

EXHIBIT 6-A

Process Evaluation to Assess Integrated Services for Children

Many analysts have observed that the traditional system of categorical funding for children's services, with funds allocated to respond to specific problems under strict rules regarding eligibility and expenditures, has not served children's needs well. The critics argue that this system fragments services and inhibits collaboration between programs that might otherwise lead to more effective services. In 1991, the Robert Wood Johnson Foundation launched the Child Health Initiative to test the feasibility of achieving systemic changes through the integration of children's services and finances. Specifically, the initiative called for the development of the following components:

- A decategorization mechanism that would pool existing categorical program funds and create a single children's health fund
- A care coordination procedure using case management that would use the pooled funds to provide comprehensive and continuous care for needy children
- A monitoring system that would identify the health and related needs of children in the community and the gaps in existing services

Nine sites across the country were selected to launch demonstration programs. The Institute for Health Policy Studies, University of California, San Francisco, conducted an evaluation of these programs with two major goals: (1) to gauge the degree to which the implementation of the projects was consistent with the original planning objectives (fidelity to the model) and (2) to assess the extent to which each of the major program components was implemented. In the first year, the evaluation focused on the political, organizational, and design phase of program development. During subsequent years, the focus turned to implementation and preliminary outcomes. A combination of methods was used, including site visits, written surveys completed by the program managers, in-depth interviews of key participants, focus groups of service providers and clients, and reviews of project-related documents.

The evaluation found that most of the nine sites experienced some degree of success in implementing the monitoring and care coordination components, but none was able to implement decategorization. The general findings for each component were as follows:

- Decategorization: Several sites successfully created small pools of flexible funds, but these were from sources other than categorical program funds. No site was able to fully implement decategorization under the definitions originally adopted.
- Care coordination: This was implemented successfully by most of the sites at the client level through case management, but there was generally less coordination at the system level.
- Monitoring: The sites encountered a number of barriers in successfully completing this task, but most instituted some appropriate process.

SOURCE: Adapted from Claire Brindis, Dana C. Hughes, Neal Halfon, and Paul W. Newacheck, "The Use of Formative Evaluation to Assess Integrated Services for Children," *Evaluation & the Health Professions*, 1998, 21(1):66-90.

services. Program process is often the focus of formative evaluation designed to provide useful feedback to managers and sponsors of new programs. In the case of a more established program, a process evaluation might be called for when questions arise about how well the program is organized, the quality of its services, or the success with which it is reaching the target population. A process evaluation may also constitute the major evaluation approach to a program charged with delivering a service known or presumed effective, so that the most significant performance issue is whether that service is being delivered properly. In a managed care environment, for instance, process evaluation may be employed to assess whether the prescribed medical treatment protocols are being followed for patients in different diagnostic categories.

The second major role of process or implementation evaluation is as a complement to an impact evaluation. Indeed, it is generally not advisable to conduct an impact evaluation without including at least a minimal process evaluation. Because maintaining an operational program and delivering appropriate services on an ongoing basis are formidable challenges, it is not generally wise to take program implementation for granted. A full impact evaluation, therefore, includes a process component to determine what quality and quantity of services the program provides so that this information can be integrated with findings on what impact those services have.

Continuous Program Process Evaluation (Monitoring) and Management Information Systems

The second broad form of program process evaluation consists of continuous monitoring of indicators of selected aspects of program process. Such process monitoring can be a useful tool for facilitating effective management of social programs by providing regular feedback about how well the program is performing its critical functions. This type of feedback allows managers to take corrective action when problems arise and can also provide stakeholders with regular assessments of program performance. For these reasons, a form of process assessment is often integrated into the routine information systems of social programs so that appropriate data are obtained, compiled, and periodically summarized. In such cases, process evaluation becomes coextensive with the **management information system (MIS)** in a human service program. Exhibit 6-B describes an MIS that was developed for a marital and family counseling program.

MISs routinely provide information on a client-by-client basis about services provided, staff providing the services, diagnosis or reasons for program participation, sociodemographic data, treatments and their costs, outcome status, and so on. Some of the systems bill clients (or funders), issue payments for services, and store other information, such as a client's treatment history and current participation in other programs. MISs have become the major data source in many instances because much

EXHIBIT 6-B

An Integrated Information System for a Family and Marriage Counseling Agency in Israel

The Marital and Family Counselling Agency is run under the joint auspices of the Tel Aviv Welfare Department and the School of Social Work at Tel Aviv University. The agency provides marital and family counseling and community services for the Jewish, Muslim, and Christian residents of one of the poorest sections of Tel Aviv.

The integrated information system developed for the agency is designed to follow up clients from the moment they request help to the end of treatment. It is intended to serve the agency and the individual counselors by monitoring the process and outcomes of treatment and providing the data needed to make organizational and clinical decisions. To accomplish this, data are collected on three forms and then programmed into the computerized information system. The data elements include the following:

- Background data provided by the client, for example, sociodemographic characteristics, medical and psychological treatment history, the problems for which they are seeking help, the urgency of those problems, their expectations from treatment, and how they found out about the clinic.
- The McMaster Clinical Rating Scale, a standardized scale that monitors families on the basis of six dimensions of family functioning and overall family health; the counselors fill out this form once a month for each client.
- Retrospective evaluation forms filled out after treatment is completed, one by the counselors and another by the clients. This includes, for example, factual questions about the treatment such as its duration, the problems dealt with, the degree to which the client and counselor agreed on the problems, whether there were issues not addressed and why. It also includes retrospective assessments of the process, evaluations of improvement in the presented problems and the McMaster areas of functioning, and client and counselor satisfaction with the process and outcomes.

The counselors can enter and retrieve data from this system whenever they wish and are given a graph of each client's status every three months to support clinical decisions. Also, reports are generated for the clinic's management. For example, a report of the distribution of clients by ethnic group led to the development of a program located within Arab community centers to better reach that population. Other management reports describe the ways and times at which treatment is terminated, the problems that brought clients to the agency, and the percentage of people who applied for treatment but did not show up for the first session. The information system has also been used for research purposes. For example, studies were conducted on the predictors of treatment success, the comparative perceptions by clients and counselors of the treatment process and outcomes, and gender differences in presenting problems.

SOURCE: Adapted from Rivka Savaya, "The Potential and Utilization of an Integrated Information System at a Family and Marriage Counselling Agency in Israel," *Evaluation and Program Planning*, 1998, 21(1):11-20.

of the information that otherwise would have to be gathered in data collection for process monitoring is available in the program's MIS. Even when a program's MIS is not configured to completely fulfill the requirements of a thoroughgoing process evaluation, it may nonetheless provide a large portion of the information an evaluator needs for such purposes. MISs can thus supply data that can be used by both managers and evaluators.

Perspectives on Program Process Monitoring

There is and should be considerable overlap in the purposes of process monitoring whether they are driven by the information needs of evaluators, program managers and staff, or policymakers, sponsors, and stakeholders. Ideally, the monitoring activities undertaken as part of evaluation should meet the information needs of all these groups. In practice, however, limitations on time and resources may require giving priority to one set of information needs over another. Although there are many exceptions, the perspectives of the three key "consumer groups" on the purposes of program monitoring typically vary. These differences in perspective apply generally to outcome monitoring as well.

Process Monitoring From the Evaluator's Perspective

A number of practical considerations underlie the need for evaluation researchers to monitor program process. All too often a program's impact is sharply diminished and, indeed, sometimes reduced to zero because the appropriate intervention was not delivered, was not delivered to the right targets, or both. We believe that more program failures are due to such implementation problems than to lack of potentially effective services. Process monitoring studies, therefore, are essential to understanding and interpreting impact findings. Knowing what took place is a prerequisite for explaining or hypothesizing why a program did or did not work. Without process monitoring, the evaluator is engaged in "black box" research with no basis for deciding whether a larger dose of the program or a different means of delivering the intervention would have changed the impact results.

Process Monitoring From an Accountability Perspective

Process monitoring information is also critical for those who sponsor and fund programs. Program managers have a responsibility to inform their sponsors and funders of the activities undertaken, the degree of implementation of programs, the problems encountered, and what the future holds (see Exhibit 6-C for one perspective