

Pacific Northwest National Laboratory

Operated by Battelle for the
U.S. Department of Energy

The Science Manager's Resource Guide to Case Studies

K.M. Branch, PNNL
M.S. Peffers, PNNL
R.T. Ruegg, TIA Consulting
R.W. Vallario, U.S. DOE

September 2001

Prepared for U.S. Department of Energy, Office of Science
Office of Planning and Analysis
Under Contract DE-AC06-76RL01830



DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC06-76RLO 1830

Printed in the United States of America

**Available to DOE and DOE contractors from the
Office of Scientific and Technical Information, P.O. Box 62, Oak Ridge, TN 37831;
prices available from (615) 576-8401.**

**Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161**



This document was printed on recycled paper.

The Science Manager's Resource Guide to Case Studies

K.M. Branch, PNNL
M.S. Peffers, PNNL
R.T. Ruegg, TIA Consulting
R.W. Vallario, U.S. DOE

September 2001

Prepared for
the U.S. Department of Energy, Office of Science
Office of Planning and Analysis
Under Contract DE-AC06-76RL01830

Pacific Northwest National Laboratory
Seattle, Washington 98105

This page intentionally left blank

The Department of Energy's Office of Science is a principal steward of the physical sciences and an important contributor to the biological, environmental, and computational sciences. As with any science institution, a primary goal is the creation of knowledge – an output that is difficult to anticipate or measure under the best of circumstances. Over the last several years, the issue of science program measurement and evaluation has been a topic of great debate as agencies gain more experience in implementing the Government Performance and Results Act of 1993 and encounter, first-hand, some of the unique challenges of evaluating and documenting the performance of their science programs. Well-designed case studies can provide a rich combination of qualitative and quantitative information, offering valuable insights into the nature, outputs, and longer-term impacts of basic research, built upon objective, systematic, and credible methods.

The growing attention to accountability and performance in publicly funded science programs provides considerable impetus for research managers and administrators to increase their understanding of evaluation and assessment methods. Inevitably, such individuals are called upon to plan, commission, contribute to and/or participate in evaluative studies, whether it is to more generally communicate the benefits of research, or to present and defend programs and budgets. When the time arrives, manager's who have adequately prepared will have advantage in assuring that the evaluations are of high quality and that they are framed, interpreted and applied correctly. Similarly, analysts who typically perform case studies depend on collaborations and inputs from those who best understand the broad perspective of the scientific research, including program managers, project managers, and principal investigators. Assuring that these senior research professionals are familiar with general case study methods can improve collaborations, and ultimately, results and subsequent use of the evaluations.

Believed to be a first of its kind, this guide takes the science manager through the steps of planning, implementing, validating, communicating, and using case studies. It outlines the major methods of analysis, describing their attributes, strengths and weaknesses, and applicability while providing examples and supporting information from a tutorial perspective. Prepared for the Office of Science, the guide is designed to strengthen the Department of Energy's own processes for evaluation, although the techniques and methods are likely to be of interest to science managers everywhere. The current version may be viewed as a work in progress and updates will follow as more experience and feedback are obtained. Your comments and advice are welcomed.

Bob Vallario

Office of Science
US Department of Energy

ACKNOWLEDGEMENTS

The authors would like to thank all those who contributed to the content of this resource guide. This includes the many staff members at DOE and PNNL who expressed interest in and ideas for this guide, particularly those who attended the Workshop on Guidance for a DOE-SC Case Study Methodology held at the Battelle Washington, D.C. Office on December 12, 2000. Specific thanks go to Francis Narin (CHI Research, Inc.), Irwin Feller (Penn State), Joan Adams (PNNL), Charlette Geffen (PNNL), Elizabeth Malone (PNNL), Jim Dooley (PNNL), Bill Valdez (DOE), and Rick Borchelt (DOE) for their contributions and support.

TABLE OF CONTENTS

Preface	i
Acknowledgements	ii
Table of Contents	iii
1. Introduction	1
1.1. Case Studies: A Research Tool for Understanding Programs in Their Real-Life Context	1
1.2. Why Conduct Case Studies?	2
1.2 Organization of the Guide	2
2. Case Study Fundamentals	5
2.1 The Attributes of Good Case Studies	5
2.2 Benefits of Multiple Methods and Both Qualitative and Quantitative Information	7
2.3 Placing the Case Study within the Office of Science Context	7
3. Planning a Case Study	11
3.1 General Approach	11
3.2 Establish the Purpose and Scope of the Study	12
3.3 Identify and Characterize the Audience	13
3.4 Determine the Key Questions	14
3.5 Select the Appropriate Time Frame for the Study	14
3.6 Choose an Appropriate Case or Set of Cases	17
3.7 Establish Collaborations and Partnerships, If Appropriate	17
3.8 Select Appropriate Data Collection and Analytic Methods	18
3.9 Match Resources and Schedules with Methods and Approach	19
3.10 Develop a Preliminary Study Design	25
3.11 Conduct an Initial Internal and External Review	25
3.12 Commission the Analysts and Finalize the Study Design	26
4. Managing Case Study Implementation	29
4.1 Provide Management Updates	29
4.2 Monitor Common Data Collection and Analytic Difficulties	29
4.3 Maintain Coordination Among Collaborators and Partners	30
4.4 Track Key Issues in Case Study Data Analysis	31
4.5 Oversee Preparation of the Case Study Report and Oral Presentation Materials	32
4.6 Coordinate the Review, Revision, and Finalization of Study Report(s)	33
5. Communicating and Applying Case Study Results	35
5.1 Gauge the Target and Plan the End Game	35
5.2 Strengthen the Message: Build On Case Study Findings	36
5.3 Share Findings with Secondary and Tertiary Audiences	36
6. References	39
Appendix A. Eleven Case Study Methods: Overviews and Supporting Materials	1

A.1 Classic, Interview- and Records-Based Method.....	A-1
A.2 Expert Review Method	A-5
A.3 Process Outcomes Method	A-9
A.4 Historical Tracing Method.....	A-13
A.5 Bibliometrics and Citation Methods	A-17
A.6 Content Analysis Methods.....	A-23
A.7 Sociometric / Social Network Methods	A-27
A.8 Survey Method, Including User and Participant Surveys.....	A-31
A.9 Benefit-Cost (Cost-Benefit) Methods	A-35
A.10 Statistical/Econometric Methods	A-41
A.11 Key Indicator Method.....	A-45

TABLE OF TABLES

Table 1. Organizational Themes Historically Important to the Office of Science	8
Table 2. Key Components of Evaluation and Performance Measurement	9
Table 3: Characterizing Each Audience for the Study	15
Table 4. Questions that Office of Science Case Studies Might Address.....	16
Table 5. Factors to Consider in Selecting the Cases to Study.....	18
Table 6. Overview of Candidate Case Study Methods	20
Table 7. Matching Methods and Study Objectives.....	24
Table 8. Aspects of Data Collection Typically Causing the Greatest Difficulty.....	30

TABLE OF FIGURES

Figure 1. Case Studies Can Address the Complex Web of Research	1
Figure 2. Organization of the Guide.....	3
Figure 3. Elements of Case Study Planning	12

1.1. Case Studies: A Research Tool for Understanding Programs in Their Real-Life Context

The Office of Science defines case studies as in-depth inquiries that detail a research program, project, or facility in its real-life context. Case studies typically focus on one or a few specific examples, or *cases*, each of which is described individually.¹ Case studies typically use multiple sources of information to *describe* the case and the context in which it occurs, and may use a variety of analytical methods and tools to *identify* and *explain* particular aspects of the case. In many case studies, a key objective is to examine the links between the context, the inputs, the process, and the results to show how and why the case worked and performed as it did. Case studies are frequently used as the basis for theory development. Multiple-case studies are used to illuminate causal relationships and patterns that would otherwise be difficult to see.

Figure 1 illustrates these complex causal relationships and patterns. This complexity underlies the wide range of questions Office of Science case studies address. Office of Science case studies may focus on the research process itself, one or several of the many different outputs of publicly funded research (the discoveries and knowledge, scientific networks, advanced instrumentation, new processes and procedures, and training and education), or the more distant outcomes and impacts on education, public health, economic strength, national security, environmental quality, and energy security that result from Office of Science funded research.

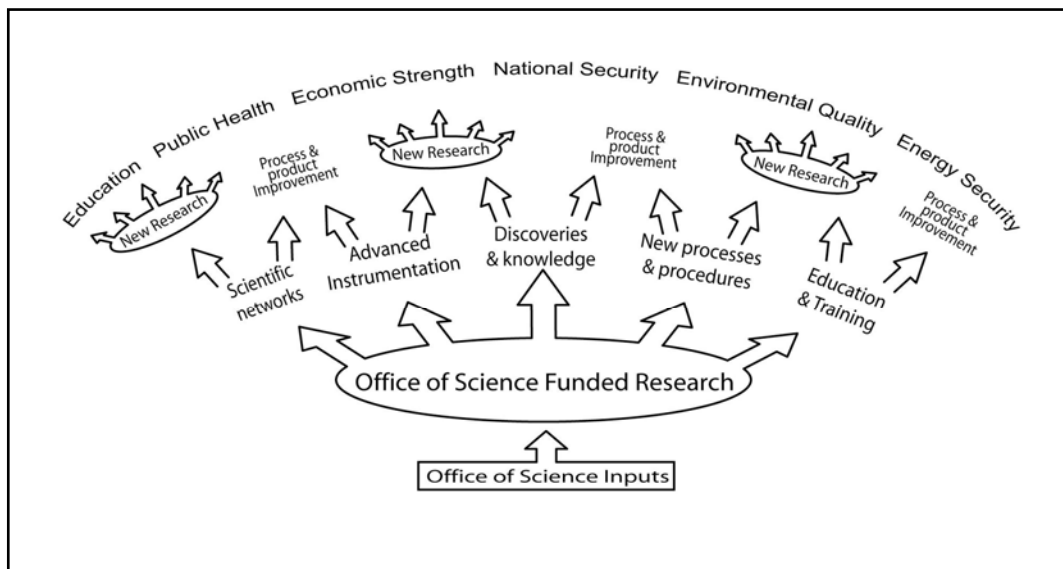


Figure 1. Case Studies Can Address the Complex Web of Research

¹ In contrast to other types of studies that aggregate cases into studies of collectives. For a good discussion of the different ways cases are addressed and aggregated in research, see Ragin and Becker's (1992) *What is a Case?*

Case studies vary widely in form and in the research methods used to assemble and analyze the information. As an example, a single case study might begin with interviews of program managers and researchers together with a review of planning documents to compile descriptive, qualitative information about the case. Analysts might then gather quantitative data on program outputs, such as publications, from an existing database. A citation analysis might be used to determine how and by whom the program's published research results are being used. A survey of users might collect data on the value of the information to the users. An econometric comparison of users and nonusers, based on publicly available data, might reveal interesting differences that assist analysts not only in assessing the potential value of the information, but also in formulating better knowledge dissemination strategies by the agency. Using multiple data sources, both qualitative and quantitative information, and multiple analysis techniques and measures, case studies can address a variety of topics of interest to Office of Science program managers.

1.2. Why Conduct Case Studies?

Office of Science programs are focused on scientific advancement and knowledge generation. They contribute to the development of fundamental science and the core missions of the Department of Energy. The Office of Science conducts case studies to describe its programs, explore particular aspects of the scientific process, and to answer questions asked about its programs by managers, participants, and other stakeholders. Case studies complement other types of research and evaluation, such as cross-sectional and topical studies and evaluations.

The case studies conducted by the Office of Science vary widely in purpose and scope. Primary objectives are to:

- ◆ *Gain management insights*, for example by assessing the quality of research, identifying pathways of knowledge flows, identifying trends and emerging areas in research, assessing program performance, and understanding the relationships among inputs, program mechanisms, outputs, and outcomes
- ◆ *Understand issues and challenges*, for example by evaluating whether collaborative relationships are working, identifying factors that influence success, determining who is using program research results and for what purpose, and examining the relationship among research areas
- ◆ *Provide inputs for strategic planning and changes in resource allocation*, for example by identifying program components that are addressing key program goals and outcomes
- ◆ *Assess core competencies and unique contributions*, for example by documenting to what extent and how past investments have paid off, assessing user satisfaction with special facilities, and describing the spheres of influence of DOE researchers
- ◆ *Communicate and explain research support*, for example by responding to specific requests for information about particular programs, projects, facilities, or capabilities, informing and educating a larger audience, and generating and presenting program metrics.

1.2 Organization of the Guide

This guide is organized to aid program managers who are considering the use of case studies and evaluators who are interested in conducting case studies of science projects and programs –

particularly those funded by governmental agencies. It outlines the steps involved in managing case studies of scientific research programs. It also provides brief descriptions of the major data collection and analytic methods used in case studies, a summary of their relative merits, and references to sources that provide more detail.

Case study management involves planning, implementation, validation, communication, and application of case study research. Science managers who understand this process are in a stronger position to conduct studies of high quality and to use the results accurately and effectively. They can collaborate more effectively with case study analysts to design studies that provide the best combination of qualitative and quantitative information to address important and meaningful questions.

👍 Successful case studies are built equally upon knowledge of the case study methods and subject matter.

Figure 2 illustrates the outline of the guide. Following this introduction, Chapter 2 describes case study fundamentals. Chapters 3 and 4 cover the planning and implementation phases of case study research, respectively. Chapter 5 provides suggestions for communicating and using case study results. Appendix A presents an overview and supporting materials for eleven candidate data collection and analytic methods that might be used in Office of Science case studies.

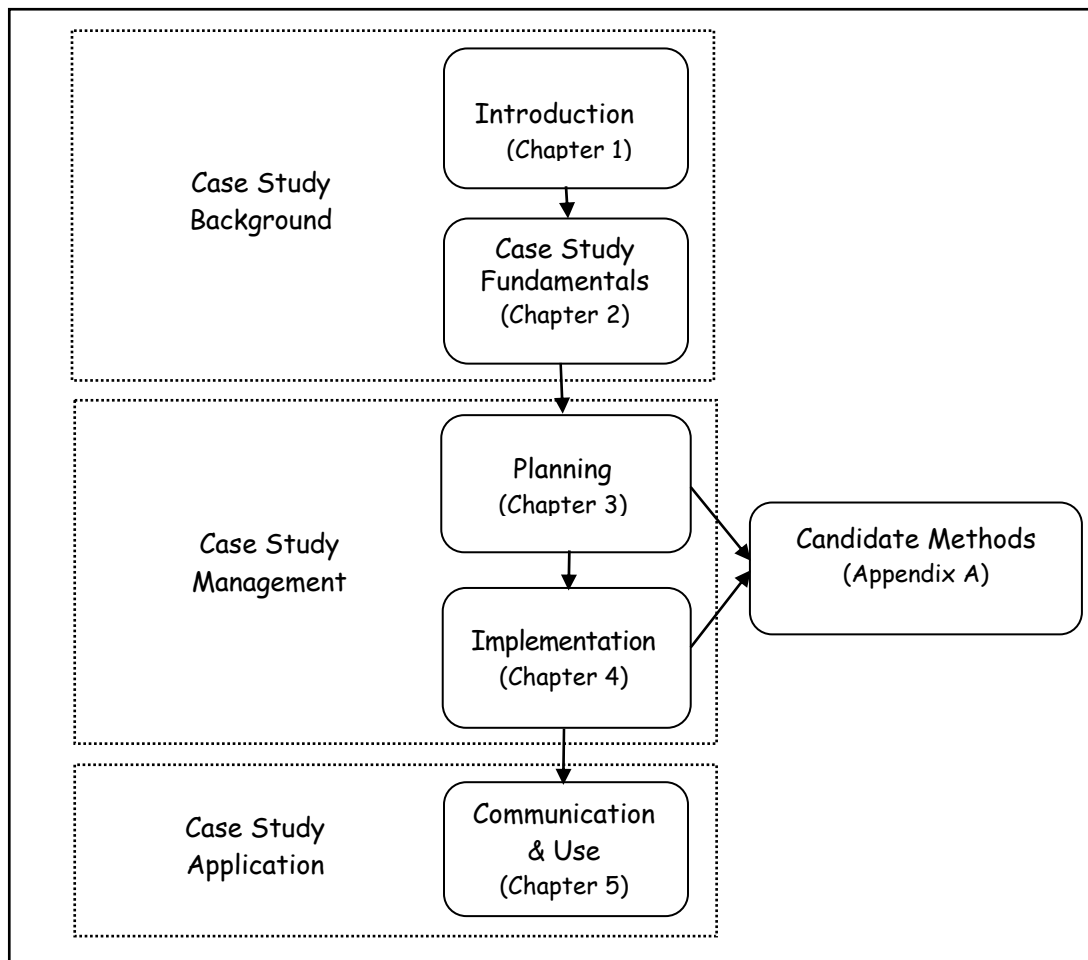


Figure 2. Organization of the Guide

This page intentionally left blank

2. CASE STUDY FUNDAMENTALS

2.1 The Attributes of Good Case Studies

Case studies range from short, focused descriptions of projects or programs to extensive, book-length accounts of emerging fields.² Whether long or short, case studies are a form of research and they are expected to conform to the standards of good research. There is general agreement that good, rigorous case studies address significant questions and hypotheses that are substantive, subject to dispute, and specific (Booth et al. 1995:94-96). Good case studies employ sound research practices and support their claims with reliable and valid evidence (Booth et al. 1995:96-104; Yin 1984:140-145).

Reliable evidence is:

- ◆ Relevant
- ◆ Accurate
- ◆ Confirmable
- ◆ Representative
- ◆ Appropriate
- ◆ Sufficient
- ◆ Authoritative

Case study research is conducted through observations and analyses rather than experiments. Frequently, the boundaries between the case and the context are not clearly evident (Yin 1984:23). Unlike laboratory experiments, which may allow for “all else to be held constant” except the object of study, case studies occur in a context too complex to allow for this type of controlled experimentation. Consequently, case study analysts must link their evidence to their claims through logical steps based on sound principles and rigorous analysis, with a goal of establishing high internal and external validity. As discussed by Booth et al. (1995:112) in *The Craft of Research* and by Barzun and Graff (1992) in *The Modern Researcher*, this is accomplished by careful design, use of multiple methods, and attention to competing evidence. In general, examining multiple, strategically selected cases strengthens the study and the credibility of its conclusions. Good case studies, like good experiments, provide a solid basis for analytical generalization. That is, they generalize the particular findings of the case study to broader theories of science performance and social impact (Yin 1984:39).

Good case studies consider alternative perspectives and explicitly develop and evaluate alternative, competing explanations for their key observations. Case studies may use control groups and counterfactual scenarios to strengthen claims of attribution and impact. These formal approaches add credibility to the answers case studies provide to two questions frequently asked of managers of publicly funded research: “What value resulted from the public funding?” and “What would have happened without the public funding?” Consequently, although it can be difficult to identify appropriate control groups or counterfactuals for some aspects of basic research, it is often worth considerable effort to gain this additional credibility.

👍 Good case studies consider alternative explanations for key findings.

² Examples of short, focused descriptive case studies are those used by the Harvard Business School; an example of an extensive, book-length treatment of an emerging field is Gleick’s (1987) examination of the evolution of chaos theory.

- ◆ *Control Groups:* One approach for attributing specific results to a program is to provide comparison with an appropriate control group. For example, in the case of federal technology partnership programs, control groups have been constructed from companies that did not receive a federal award or grant, for comparison with those that did. Though it is often not possible to create a scientifically designed, randomly assigned control group for Office of Science programs, **it may be possible to approximate a similar group or situation for purposes of comparison.** Use of control groups generally is a preferred way of accounting for what would otherwise have happened, if it is feasible to construct a group that offers a reasonable comparison for the topics of interest in the case study.³
- ◆ *Counterfactual Scenarios:* Another approach, usually less costly than including control groups in the study, is to develop one or more scenarios describing what might plausibly have occurred in the absence of the federal research. For example, experts in the field may have the opinion that without the federal research effort, nothing comparable to the research outputs would have been developed in the foreseeable future. This situation provides the simplest case for evaluation. It allows any benefits from federal research outputs to be attributed completely to the federal program. Or the experts may estimate that without the federal effort comparable research outputs would have been delayed by some specified number of years. The analyst can then compute the value of having the research outputs accelerated by the specified number of years. Such scenarios may address other ways federally funded research affects outcomes by, for example, changing the scale or scope of research in a field or changing the probability of success of subsequent research.⁴

At a minimum, good case studies consider an analytic “future without” (or, for completed programs, a “present without”) the program or activity being studied. Questions to ask include: Would some other organization likely have done something similar? Who? What? In what time frame? What would users of the program’s outputs have done without them? What difference would it have made if they had used the next best substitute? This line of thinking can help put the significance of case study findings in perspective when no detailed counterfactual or control-group analysis has been performed.

Finally, case studies are also a form of communication, and should conform to the standards of effective communication. Good case studies are clear and engaging. They tell a compelling story and provide the basis for action, where appropriate. As Gregory and Miller (1998:138) observe: “Scientific research makes a good story if it can be followed both as science and as narrative at the same time.” *The Craft Of Research* (Booth et al. 1995) provides excellent guidance on the development of effective research reports.


³ For example of studies that feature control groups, see Feldman and Kelley (2001) and Watkins (forthcoming). The Science Policy Research Unit (1972) used pair wise comparisons of successful and unsuccessful innovations to provide some of the benefits of control groups. Note that these examples are from technology programs not basic science programs, where the approach can be more difficult to apply.

⁴ For an example of a counterfactual analysis based on expert opinion, see Martin et al. (1998). David et al. (1992) discuss the limitations of counterfactuals in assessing the economic payoffs of basic research.

2.2 Benefits of Multiple Methods and Both Qualitative and Quantitative Information

The Office of Science is frequently called upon to describe, explain, and document the value it adds to the state of science. Quantitative measures alone often can't address the nature or significance of these achievements or the complex relationship of the science to the broader context in which it occurs. Likewise, qualitative descriptions alone may lack the rigor and clarity to be convincing.

The best case studies use multiple methods and qualitative and quantitative data drawn from multiple sources to create a more complete picture. Comprehensive descriptive information not only makes the case more interesting, persuasive, and memorable to the audience, it often provides the best basis for analyzing the case. It helps identify and examine hypothesized patterns and cause-and-effect relationships. It also helps characterize attributes important to the Office of Science and its stakeholders, such as significance, influence, leadership, and value. Quantitative data allows the analyst to present summary measures and add statistical rigor.

 The best case studies use multiple methods and both qualitative and quantitative data.

Case studies draw upon the full range of data sources and data collection methods used in the social sciences, including:

- ◆ Documents and archival records
- ◆ Interviews, focus groups, and surveys
- ◆ Direct observations
- ◆ Physical artifacts.

They also draw on the full range of analytical methods, including those highlighted in Appendix A. The value of using multiple methods, sources of information, and measures is that the resulting case study is more descriptive, robust, and defensible, in part because the multiple sources of information are used to triangulate the results, thus increasing confidence in the findings.

2.3 Placing the Case Study within the Office of Science Context

Office of Science case studies should reflect the organization's key goals and themes in the framing of questions, the formulation of reports and presentations, and the application of results. Why? Because in addition to meeting the primary purpose of the study, the case study manager should make every reasonable effort to relate the study to broader organizational goals and themes. This serves two functions.

First, it establishes the broader relevancy of the case study and, ultimately, the subject matter being addressed. Second, it provides a framework for linking the independent case studies through common messages. This structure and the corresponding linkages become powerful communication tools as the whole becomes greater than the sum of the parts. Case studies can also be used to test, validate, and, if appropriate, challenge the organization's themes.


 Relating the case study to the organization's broad themes establishes relevancy.

Table 1 presents organizational themes that have been important to the Office of Science. Organizations and organizational emphasis change with time, as do the ways that an organization

Table 1. Organizational Themes Historically Important to the Office of Science

- ◆ *The nation's principal steward and largest supporter of the physical sciences, building foundations for energy options and knowledge for the sciences.*
- ◆ *Significant contributors to the biological, environmental, and computational sciences, with expertise built around our mission and core competencies.*
- ◆ *Source of the scientific knowledge and discoveries that underpin DOE's applied missions in energy, environmental quality, and national security, creating new options and pathways to success.*
- ◆ *An organization whose primary "product" is the creation of knowledge, whether through our basic research or our support for the critical infrastructure that empowers our nation's scientific community.*
- ◆ *A provider to general science community of many of the largest, most powerful research instruments at the cutting-edge of science, award-winning builders and operators of the definitive probes to explore matter, energy, and life itself.*
- ◆ *A funder of research that is so fundamental and foundational that it has a strong multiplier effect on many fields of science, as well as the research efforts of other leading research institutions.*
- ◆ *A leading incubator for interdisciplinary research, exploring some of the more complex, stubborn scientific mysteries and forging linkages between universities and national laboratories spanning many synergistic disciplines.*
- ◆ *A sponsor of science that spans scales from the infinitesimally small, to the infinitely large, building linkages through the continuum.*
- ◆ *A team player on many scientific investigations, partnering on many national and international science challenges.*
- ◆ *An effective program manager whose programs and directions are driven largely by the scientific community, building on an extensive advisory committee process and a highly credentialed, largely scientific federal project staff.*
- ◆ *A sponsor of a balanced research portfolio, split between our nation's universities and our national laboratories.*
- ◆ *An informed leader, responsive to, but not driven by the research needs of the applied research community, assuring that our research selection and planning is informed by potential applications.*
- ◆ *A supporter of education, affording unique opportunities in graduate-level math and science education and providing for the nation's next generation of world-class scientists.*
- ◆ *A cornerstone of the basic science enterprise of the United States, with a history of discovery as impressive as it is far-reaching.*

chooses to communicate its major features and strengths. Consequently, the case study's sponsors and planners need to clarify the themes pertinent to the study. This is particularly important during periods of transition, such as during a change in administration, when key positioning themes may undergo modification.

Federal agencies, including the Office of Science, are under increasing pressure to demonstrate efficient performance and measurable social benefit. The Government Performance and Results Act (GPRA), enacted in 1993, calls for increasingly clear articulation and measurement of the performance and results of government activities and requires all federal agencies to measure and report on the results of their activities each year. Case studies often draw upon and extend the information used in these performance evaluations, and use some of the same terminology. Consequently, it is useful to be familiar with GPRA terminology, as shown in Table 2.

Table 2. Key Components of Evaluation and Performance Measurement

<p>Performance goals</p> <ul style="list-style-type: none"> ◆ Target levels of performance, expressed as tangible, measurable objectives against which actual achievement can be compared <p>Inputs (Costs), which include</p> <ul style="list-style-type: none"> ◆ Financial assets ◆ Capital assets including real estate, physical plant, equipment (including descriptive characteristics) ◆ Revenues and allocations from funding sources (e.g., for public agencies) ◆ Human resources (e.g., numbers of employees by job category, social and educational backgrounds, training, and previous experience) <p>Outputs (activities or efforts, typically seen as intermediate benefits), which include</p> <ul style="list-style-type: none"> ◆ Publications ◆ Research facilities provided ◆ Discoveries ◆ Models ◆ Human Outputs (scientists trained, collaborations established) <p>Outcomes (an assessment of the results of a program activity compared to its intended purpose), which include</p> <ul style="list-style-type: none"> ◆ Improved scientific performance (e.g., order of magnitude increase in computing speed) ◆ Scientific quality ◆ Scientific relevance ◆ Scientific leadership <p>Impacts (ultimate societal consequences), which include</p> <ul style="list-style-type: none"> ◆ Social benefits, such as improved health, increased productivity, and higher quality of life ◆ Negative outcomes, such as environmental contamination
--

This page intentionally left blank

3. PLANNING A CASE STUDY


3.1 General Approach

Planning and managing a case study is similar in many ways to planning and managing any other research effort. But the particular methods, tools, data, and study topics of case studies may be relatively new to most science managers. Since familiarity with methods and tools influences planning decisions, those planning to initiate and conduct case studies should:

- ◆ Work through the elements in the planning process, assembling available information and identifying questions that need to be answered.
- ◆ Read Appendix A to gain familiarity with the various data collection and analytic methods.
- ◆ Obtain the necessary expert assistance, and then work again through the elements in the planning process to develop a concrete plan for the case study.

This chapter is laid out as if planning were a linear process. In reality planning a case study involves a series of iterations through the following key elements:

- ◆ Establish the purpose and scope.
- ◆ Identify and characterize the audience for the study.
- ◆ Determine the key questions or hypotheses.
- ◆ Select the appropriate time frame.
- ◆ Choose an appropriate case or set of cases.
- ◆ Establish collaborations and partnerships with others interested in the case study.
- ◆ Select appropriate data collection and analytic methods (see Appendix A).
- ◆ Match resources and schedules with methods and approach.
- ◆ Develop a preliminary study design.
- ◆ Conduct an initial internal and external review.
- ◆ Commission the analysts and finalize the study design.

 Case study planning involves iteration. It is not a linear process.

These planning elements are highlighted in Figure 3, and described more fully in the remainder of the chapter.

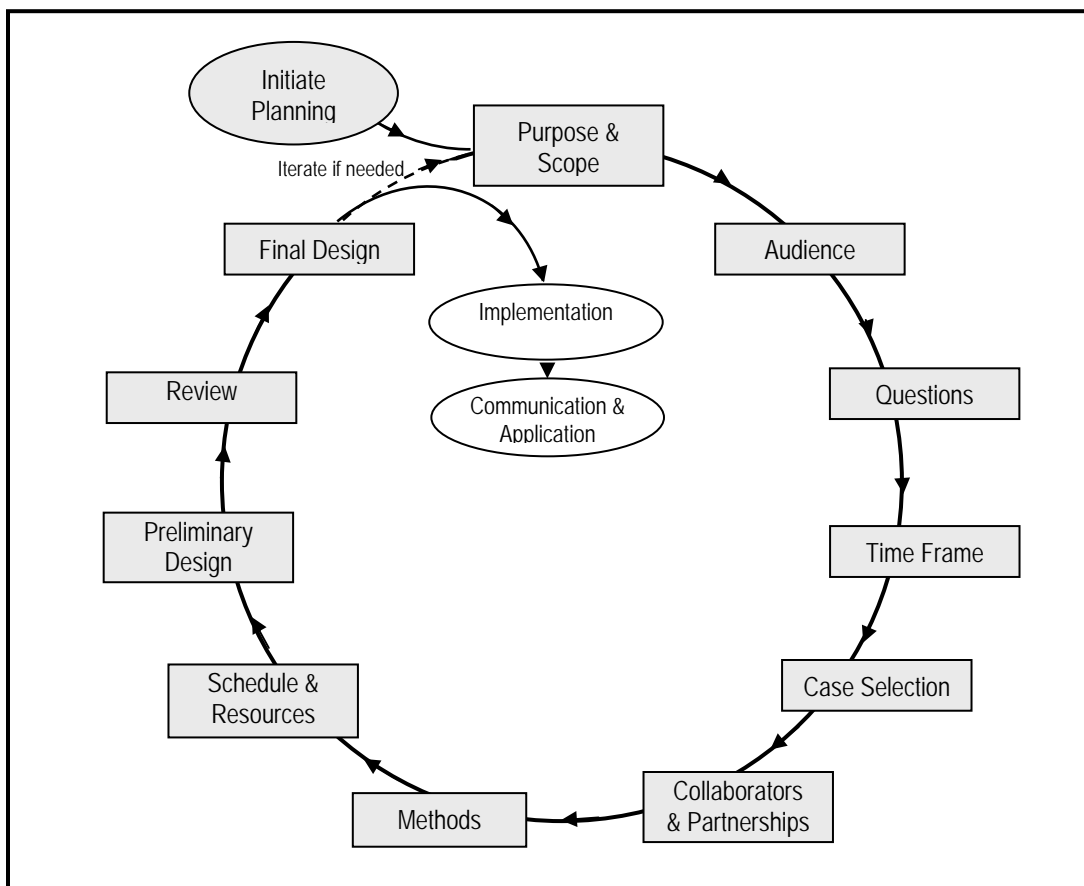


Figure 3. Elements of Case Study Planning

3.2 Establish the Purpose and Scope of the Study

Case studies in the Office of Science can be initiated in a number of ways. Congress or an external advisory committee can mandate the purpose and scope, as well as the case selection and/or schedule of a case study. Upper management or the Policy and Analysis Office sometimes provide the mandate. Sometimes the impetus for a case study arises at the program level, leaving decisions about purpose, scope, and case selection up to the program manager. Regardless of how the process is initiated, the program manager needs to clearly and quickly establish why and for whom the study is to be done, and what expectations or constraints will affect its scope, case selection, focus, and approach.

👉 Purpose and scope need to be aligned with resource availability and schedule requirements.

As with all research enterprises, initial expectations may be unrealistic. Resource and schedule constraints may limit the research effort, requiring the scope to be scaled back and the focus narrowed. Matching the case study to resources and schedule restraints is part of the planning process (see section 3.9 for more information).


Before focusing on the particulars of design, it helps to assess the “lay of the land” by asking a series of key questions.

Key questions to get the "lay of the land:"

- ◆ Why are we doing this case study?
- ◆ Who is the audience?
- ◆ Are the results needed by a particular time? If so, when?
- ◆ What kind of information are we looking for?
- ◆ What kind of data and resources are available? Where?
- ◆ What positioning issues are there?
- ◆ Which program/project/facility/process/capability/discovery are we focusing on? Why?
- ◆ Who should we work and partner with to perform the case study?
- ◆ What data collection and analytic methods are most appropriate?

3.3 Identify and Characterize the Audience

The interests and concerns of the audience drive the purpose and scope of the study as well as the structure and style of the research report. Different audiences have different interests. Good case study designs respond to multiple audience needs, placing a priority on addressing the interests of the primary audiences. The audiences' questions focus the study and influence the selection of the cases and methods.

 Good case study designs respond to multiple audience needs.

The primary audiences for Office of Science case studies are likely to be:

- ◆ Policy makers interested in ensuring the excellence of science and in seeing that funds are yielding important scientific contributions and leadership in particular scientific arenas
- ◆ Managers who evaluate and direct the studies to provide themselves with the insights they need to run their operations effectively. They must communicate what they are doing and why it is worth doing to a variety of audiences, and also gather evaluative information to meet organizational and legislated requirements for program metrics.


Secondary audiences for Office of Science case studies include:

- ◆ Professionals in the field of science under study interested in the technical findings of the study, the nature of the claims and the evidence presented in the study, the basis for linking the evidence with claims, and the implications. They will likely also be interested in the theory underlying the study, the quality of the research, and the validity of the findings
- ◆ Historians and social scientists interested in the detailed descriptions of science activity within a context and retrospective examination of the influence of science on society
- ◆ The science-attentive public, who focus primarily on knowledge creation and the process of scientific discovery
- ◆ The broader public, who are interested in the relevancy and social benefits of the science.

Tertiary audiences for Office of Science case studies include:


- ◆ Professionals in the field of evaluation, who are interested in the research design, methods, and results
- ◆ Managers in other organizations, who are interested in effective ways to understand program impact and measure performance.

Since most case studies will have multiple audiences, it is important to decide whose needs will be given priority. Sometimes the primary audience(s) for the study and the principal interests or concerns will be clear at the beginning of the planning process. In any event, the primary audience(s) will need to be identified and the principal interests and questions clarified as part of the planning process.⁵ The same is true for the study's secondary audiences. Find out what questions the audiences have, and what type of evidence they would find most convincing. This influences both the study design and the presentation of results. Booth et al. (1995:26-27) recommend using a checklist similar to that shown in Table 3 to make sure each target audience is understood. Check back through the list as the study progresses to keep this information in mind.

 Keep your target audience's needs and interests in mind throughout the study.

3.4 Determine the Key Questions

The key questions the audiences pose drive both the selection of cases and the design and methods of the study. It is important to develop a clear understanding of the questions the audience wants the case study to answer. If the case study is being initiated in response to a specific mandate or request (for example, from Congress), the key questions may already be well articulated and relatively immutable. If the case study is being initiated internally, there is more opportunity to formulate the key research questions that will guide the case study design.

 Insightful questions are the key to valuable research.

Case studies can address a wide range of questions. Insightful questions matched with appropriate methods lead to valuable research. Table 4 and Appendix A illustrate the types of questions case studies can address.

3.5 Select the Appropriate Time Frame for the Study

Case studies differ from other forms of evaluation or research. They focus on a particular event or set of events with attention to their placement within a broader, real-life context. Although they may sometimes include projections into the future,⁶ case studies typically address cases that either:

- ◆ Occurred in the past and are completed (retrospective or ex-post studies)
- ◆ Started in the past and are ongoing (ongoing or contemporary studies)
- ◆ Start concurrently with the study (monitoring studies or concurrent studies).

⁵ This may involve meeting with or interviewing members of the potential audience.

⁶ Studies focusing primarily on anticipated benefits and impacts are generally categorized by the Office of Science as “feasibility studies,” even though they use many of the same methods and have many of the same characteristics as case studies.

Table 3: Characterizing Each Audience for the Study

Questions to answer about each audience:

What do they expect the study to do for them?

- ◆ Provide a basis for making a particular decision or taking a particular action?
- ◆ Assess a decision that has been made in the past?
- ◆ Help them solve a current problem?
- ◆ Help them understand something better?

How much do they already know?

- ◆ More, less, or about the same knowledge of a) the background area; b) the particular topic; and c) the methods of analysis as those conducting the study?
- ◆ Do they have a special interest in the topic of the study?
- ◆ Are there particular aspects of the topic they expect the study to address?
- ◆ What evidence will they be looking for, or have particular confidence in?

Do they already understand the problem/question?

- ◆ Did they pose the problem/question?
- ◆ Do they share the problem/question but have not yet recognized it?
- ◆ Do they understand the problem/question, even if it is not within their purview?
- ◆ Do they have an alternative formulation of the problem/question?
- ◆ Do they think that the problem/question is important and worthy of study?

How will they respond to the solution/answer?

- ◆ Do they expect to act based on the results of the study?
- ◆ Will the results confirm or contradict what they already believe?
- ◆ Will they be surprised by the findings or the nature of the evidence?
- ◆ Will they challenge the results and require extensive evidence and justification of the methods?

In what forum will they encounter the report?

- ◆ Have they asked for the report or will it be presented to them unbidden?
- ◆ Before it reaches the main readers, will the report have to be approved by gatekeepers or intermediaries?
- ◆ Is there an established format they will expect the report to follow?

Adapted from Booth et al. 1995:26-27.

Table 4. Questions that Office of Science Case Studies Might Address

Case studies might be used to answer questions about:

Management Insights

- ◆ The quality of research
- ◆ The pathways through which knowledge generated by a program flows into the broader scientific community and beyond
- ◆ The types of influence federal research programs have on public goals and how this influence is exerted
- ◆ The trends and emerging areas within the scientific arenas being funded by a program, studied at a facility, or undertaken using particular capabilities
- ◆ The effectiveness and efficiency of program management
- ◆ The relationships among inputs, program mechanisms, and outputs

Issues and Challenges

- ◆ How well collaborative relationships and strategies are working
- ◆ What influences success in the research process
- ◆ Who is using research results and for what
- ◆ The relationships among research areas


Core Competencies and Unique Contributions

- ◆ The payoff from past investments
- ◆ User satisfaction with special facilities
- ◆ Spheres of influence of researchers

Communicating and Explaining Research

- ◆ The specifics of particular programs, projects, facilities, or capabilities
- ◆ How to interest/educate/inform a larger audience
- ◆ Program performance and metrics
- ◆ How particular research led to new discoveries, theories, or insights

Studies attempting to answer some key questions for the Office of Science – for instance on the impact of its discoveries – often need to cover a relatively long time period. Generally, the more basic the research, the longer the elapsed time before program outcomes and ultimate impacts occur. Technology programs generally show impact more quickly than basic research programs.


 The more basic the research being studied, the longer elapsed time is needed to capture impacts.

The timing and duration of the study can significantly affect the type of information that is available. People with knowledge of the case may become busy with other topics, move, retire, forget, and die. If not collected relatively quickly, this knowledge may become lost or distorted. Thus it is very important to match the timing and time frame of a study with the questions to be addressed, the case(s) to be selected, and the methods to be used. Timing considerations are

discussed in the descriptions of methods in Appendix A, with special attention in the discussion of Cost-Benefit Methods (A.10), where timing is particularly important to the capture and computation of ultimate, cumulative benefits.

3.6 Choose an Appropriate Case or Set of Cases

As with any research, selecting the case or cases to be studied helps determine what can be learned and which methods are most appropriate (Yin 1984; Ragin and Becker 1992). For Office of Science case studies, the unit of study, or case can vary. Bounding the scope of the case is often an important planning decision.

 Choose cases with care. Case selection can determine what can be learned from the study.

Examples of case study units:


- | | | |
|---------------|----------------|----------------|
| ◆ Programs | ◆ Sub-programs | ◆ Projects |
| ◆ Facilities | ◆ Processes | ◆ Capabilities |
| ◆ Discoveries | | |

Case selection should be treated as part of the study design if the case has not been predetermined. The case(s) should be chosen intentionally to maximize the study's ability to answer the questions and obtain the information important to the key audience(s). Table 5 summarizes some of the factors that may be most pertinent for the selection of cases for Office of Science studies.

3.7 Establish Collaborations and Partnerships, If Appropriate

The Office of Science frequently co-funds research programs, for example with US Environmental Protection Agency and National Science Foundation. Collaborating or partnering with these and other organizations can provide significant benefits, including better cooperation, facilitated access to key people and information, and improved political or financial support.

Successful collaboration requires up-front planning, an ongoing commitment to communication and coordination, and a problem-solving approach. Involving potential partners early in the planning process is key to setting a collaborative tone. It is important to establish clear ground rules regarding how findings will be presented to the participants, the stakeholders, and the public. Collaborators should have an opportunity to review findings before they are disseminated publicly. Once the ground rules are established, all participants and partners should be made aware of these protocols and the manager's commitment to them.

 Successful collaborations require planning and lots of communication.

By becoming involved early, collaborators have more opportunity to gain internal support from their own organizations and to decide how they will support the study. Early involvement enables them to participate more effectively – identifying questions they would like to have answered, identifying their audiences for the study, and participating in decisions about time frame, case selection, and methods. Early involvement is also likely to enhance the study's visibility.

Table 5. Factors to Consider in Selecting the Cases to Study

<p>Interest: Does the case provide the basis for exploring important and timely questions?</p> <ul style="list-style-type: none">◆ What has happened?◆ How it has happened?◆ Why it has happened? <p>Is the case unique (or typical) in ways that serve the research purposes?</p> <ul style="list-style-type: none">◆ Is it an exemplar?◆ Is it representative of an important problem?◆ Is it one in a planned series? <p>Access and availability: Can the necessary information be obtained?</p> <ul style="list-style-type: none">◆ Is the site and are the participants accessible and are logistics feasible?◆ Are necessary documents and people available?◆ Are key participants and other knowledgeable persons willing to participate/ cooperate? <p>Has enough happened/time passed to provide good and valid data?</p> <ul style="list-style-type: none">◆ Have the expected results had time to occur?◆ Has so much time passed that it will not be possible to trace key linkages? <p>Instrumentality: Will the findings be useful for intervention?</p> <ul style="list-style-type: none">◆ Do those involved in the study or receiving the results have control over behavioral events?◆ Will the study results be available in time to inform decisions and actions? <p>Will it provide data that contributes to a larger research agenda?</p> <ul style="list-style-type: none">◆ Does it provide a new or different type of case?◆ Does it provide an example of a different context?◆ Does it provide an opportunity to examine different outcomes/outputs?◆ Does it contribute to analysis of a portfolio of projects, programs, or capabilities?

3.8 Select Appropriate Data Collection and Analytic Methods

The methods and tools best used when conducting a case study are determined by the nature of the events being studied, the questions to be answered, and the nature of the available data. The particular capabilities, experience, resources, and preferences of those designing the study also influence method selection.

The tools and methods used in case studies have been developed primarily by social scientists. Although some are broadly applicable, many are tailored to answer particular types of questions. Consequently, they often use particular types and sources of information. Eleven methods useful for case study research are addressed in this guide.

Eleven methods useful for case study research:

- ◆ Classic, interview- and records-based method
- ◆ Expert review method
- ◆ Process-outcomes method
- ◆ Historical tracing method
- ◆ Bibliometric and citation methods
- ◆ Content analysis methods (including co-word, data tomography, data mining, and visual representations of content)
- ◆ Sociometric/social network methods
- ◆ Survey method (including user and participant surveys)
- ◆ Benefit-cost (cost-benefit) methods
- ◆ Statistical/econometric methods
- ◆ Key indicator method

Appendix A describes each method, highlighting its typical objectives, the sources of data, methods of data collection, and distinctive units of measure. Table 6 summarizes typical questions, distinctive data, units of measure, and analytic tools used by each of the eleven candidate methods. Table 7 identifies the research method(s) that might be used to address each of the primary objectives case study research for the Office of Science.

3.9 Match Resources and Schedules with Methods and Approach

The administrative aspects of planning and budgeting a case study are no different than for any other research initiative, and should pose no novel obstacles. The scope and approach of a case study must be matched with the available resources, schedule, and constraints. Success is jeopardized if the scale of the study is too great for the resources available, or if the analysts are not sufficiently skilled in the methods needed to undertake a particular study. Appendix A provides general information about the time, cost, and staff resource requirements of the eleven methods. As the parameters of the case study are established, program managers may want to consult with analysts trained and experienced in the methods of interest. Such assistance could come from analysts within DOE, a National Laboratory, or a contractor. Large, expensive studies may need to be planned with at least a year's lead-time to accommodate the governmental budget cycle.

References and resource links are provided throughout this report, particularly in Appendix A, to help identify resource needs and to identify some of the experts. When considering the staffing of evaluative studies, both skills and objectivity need to be considered. When independence and credibility are issues, consider using analysts with specialized skills and no ties to the Office of Science.

Table 6. Overview of Candidate Case Study Methods

Method	Typical Questions about Office of Science (SC) Programs	Distinctive Data, Units of Measure, Analytic Tools
Classic, Interview and Records-Based Method (Page A-1)	<ul style="list-style-type: none"> ♦ What happened of significance? ♦ What was the sequence and "chain of events?" ♦ How did the main events/consequences come about? ♦ Who made it happen? ♦ What applications resulted from this process? ♦ How did DOE-SC provide leadership? 	<p>Data</p> <ul style="list-style-type: none"> ♦ Interviews with participants, observers, and those affected ♦ Records from the case <p>Units of Measure</p> <ul style="list-style-type: none"> ♦ Events ♦ Time lines <p>Analytic Tools</p> <ul style="list-style-type: none"> ♦ Text analysis ♦ Supplementary statistical analysis of secondary data
Expert Review Method (Page A-5)	<ul style="list-style-type: none"> ♦ Is this good science? What is the quality of the work? Does it contribute to the state-of-the art in an important way? ♦ Is it contributing to the good of the Nation in an important way? How does it rank against other efforts? ♦ Does the science being funded by SC position it as a principal steward and large supporter of this field of science? Is this field of science important and consistent with DOE's mission? ♦ Have the funds devoted to this work been used wisely and for maximum impact? 	<p>Data</p> <ul style="list-style-type: none"> ♦ Evaluations from experts ♦ Interviews with experts ♦ Documentation of case components (inputs, process (team, facilities, management, etc.), outputs <p>Units of Measure</p> <ul style="list-style-type: none"> ♦ Evaluation ratings ♦ Qualitative statements <p>Analytic Tools</p> <ul style="list-style-type: none"> ♦ Statistical/mathematical analysis of ratings ♦ Text analysis
Process-Outcome Method (Page A-9)	<ul style="list-style-type: none"> ♦ How did the process that occurred yield the outputs and outcomes that were observed? ♦ Where does the science conducted with SC funding fit in a five-step discovery/innovation process? ♦ What portion of the observed outputs and outcomes are attributable to science? 	<p>Data</p> <ul style="list-style-type: none"> ♦ Interviews with participants, observers, and those affected to specify the process and output/outcomes ♦ Records from the case <p>Units of Measure</p> <ul style="list-style-type: none"> ♦ Inputs, outputs, outcomes ♦ Clusters, indices, and indicators <p>Analytic Tools</p> <ul style="list-style-type: none"> ♦ Cluster and index analysis ♦ Econometric and other process models

Method	Typical Questions about Office of Science (SC) Programs	Distinctive Data, Units of Measure, Analytic Tools
Historical Tracing Method (Page A-13)	<ul style="list-style-type: none"> ◆ How have the ideas that came out of this research influenced subsequent thinking and discovery? ◆ What is/has been the path from innovation to output to outcome and impact? ◆ What were the key factors influencing the path? What role did SC funding and/or leadership play? ◆ What does this innovation's pathway tell us about the process of scientific innovation and discovery? ◆ How does this pathway/process compare with others? 	<p>Data</p> <ul style="list-style-type: none"> ◆ Interviews with participants, observers, and those affected ◆ Records from the case ◆ Historical documents <p>Units of Measure</p> <ul style="list-style-type: none"> ◆ Antecedent and consequent events and indicators ◆ Time lines ◆ Relationships and networks <p>Analytic Tools</p> <ul style="list-style-type: none"> ◆ Link tracing ◆ Event trees ◆ Influence analysis ◆ Citation analysis
Bibliometrics and Citation Methods (Page A-17)	<ul style="list-style-type: none"> ◆ What is the state of the research landscape? Is the SC-funded science impacting this landscape? ◆ Is SC-funded science impacting national or international research? In what particular fields? ◆ Is Intellectual Property being generated from this science? Where is this occurring? Who are the participants? ◆ What kinds of multiplier effects can be linked to this science? ◆ Are the results of this science being translated and communicated to a broader audience? 	<p>Data</p> <ul style="list-style-type: none"> ◆ Citations and citation databases (e.g., for publications, addresses, conference proceedings) ◆ Patents and patent databases <p>Units of Measure</p> <ul style="list-style-type: none"> ◆ Counts of citations/patents ◆ Quality, significance rankings ◆ Links (reciprocity measures) <p>Analytic Tools</p> <ul style="list-style-type: none"> ◆ Text analysis ◆ Database management ◆ Statistical analysis
Content Analysis Methods (Page A-23)	<ul style="list-style-type: none"> ◆ What "clusters" of research exist? Where is the emphasis/funding being placed? How has that evolved over time? ◆ What organizations are funding similar or related research? ◆ How do different participants describe and explain the process? ◆ Who among the participants in the science hold similar (different) views? What attributes are associated with these groupings? 	<p>Data</p> <ul style="list-style-type: none"> ◆ Documents of various origin ◆ Interview / meeting transcripts <p>Units of Measure</p> <ul style="list-style-type: none"> ◆ Texts/documents ◆ Events ◆ Key words, phrases, and concepts ◆ Themes and sub-themes <p>Analytic Tools</p> <ul style="list-style-type: none"> ◆ Text coding and categorization

Method	Typical Questions about Office of Science (SC) Programs	Distinctive Data, Units of Measure, Analytic Tools
Content Analysis Methods (Continued)		<ul style="list-style-type: none"> ◆ Textual data mining ◆ Co-word analysis ◆ Database tomography ◆ Visual representation of text analysis
Sociometric/ Social Network Method (Page A-27)	<ul style="list-style-type: none"> ◆ Who is contributing to the research? Where are they located? ◆ How is information being distributed and shared? ◆ Are results of the research reaching the general science community? Through what networks? ◆ How are social networks influencing the nature and quantity of multiplier effects? ◆ Is the science interdisciplinary? How have the interdisciplinary teams been established? ◆ Where have the participants gone after completing this collaboration? Who are they working with? How have their careers developed? ◆ Does the research link people from different science institutions (universities, laboratories, international research organizations)? 	<p>Data</p> <ul style="list-style-type: none"> ◆ Interviews with participants, observers, and those affected (names and relationships) ◆ Records from the case (organizational charts, lists of members) <p>Units of Measure</p> <ul style="list-style-type: none"> ◆ Personal relationship patterns ◆ Network characteristics ◆ Links and pathways <p>Analytic Tools</p> <ul style="list-style-type: none"> ◆ Network analysis software ◆ Co-nomination and co-citation analysis
Survey Method (including user and participant surveys) (Page A-31)	<ul style="list-style-type: none"> ◆ Who are the users? ◆ What are their needs? Are they changing over time? ◆ How satisfied are the users with the facility and support for users? ◆ What are the characteristics of "non-users"? Why are they not using the facility/service? ◆ What does [particular group] know about the research? What is their assessment of its importance and impact? 	<p>Data</p> <ul style="list-style-type: none"> ◆ Responses to in-person, mail, phone, or e-mail questionnaires ◆ Responses to interviews, focus groups, and other interactive processes <p>Units of Measure</p> <ul style="list-style-type: none"> ◆ Counts of responses ◆ Patterns of response by respondent attributes (comparisons within and across groups of respondents) ◆ Scales and indices <p>Analytic Tools</p> <ul style="list-style-type: none"> ◆ Statistical analysis of sample survey results (SSPS, SAS) ◆ Content analysis of responses

Method	Typical Questions about Office of Science (SC) Programs	Distinctive Data, Units of Measure, Analytic Tools
Benefit-Cost (Cost-Benefit) Methods (Page A-35)	<ul style="list-style-type: none"> ◆ What have been (are) the costs of this research? When have they occurred? ◆ What have been the benefits of this research? When have they occurred? ◆ Does the monetary value of the benefits exceed the cost? In what time frame? By how much? What is the rate of return on investment? How long before costs will be recovered? Under what assumptions? 	<p>Data</p> <ul style="list-style-type: none"> ◆ Project records and budgets ◆ Identification of benefits ◆ Valuation of benefits in dollars ◆ Databases useful for valuating benefits <p>Units of Measure</p> <ul style="list-style-type: none"> ◆ Dollars of costs ◆ Dollar value of benefits ◆ Net present value dollars ◆ Benefit to cost ratios ◆ Annual rate of return on investment ◆ Years to payback <p>Analytic Tools</p> <ul style="list-style-type: none"> ◆ Economic and financial analysis software and models ◆ Discounted cash-flow analysis
Statistical/Econometric Methods (Page A-41)	<ul style="list-style-type: none"> ◆ What role has this research and its outputs/outcomes played in the larger economy? ◆ In what sectors of the economy have the impacts occurred? ◆ What is the change in productivity? 	<p>Data</p> <ul style="list-style-type: none"> ◆ National and/or regional input-output data ◆ Expenditure records from the case (wages, purchases, etc) ◆ Surveys ◆ Economic databases <p>Units of Measure</p> <ul style="list-style-type: none"> ◆ Dollars ◆ Jobs ◆ Productivity rates <p>Analytic Tools</p> <ul style="list-style-type: none"> ◆ Statistical models and analytic software ◆ Econometric models
Key Indicators (Page A-45)	<ul style="list-style-type: none"> ◆ Is the research funded by SC advancing the creation of knowledge in particular areas? ◆ How have research inputs (such as expenditures on research facilities) and outputs (such as publications) changed over time? 	<p>Data</p> <ul style="list-style-type: none"> ◆ Agency and project records (budgets, expense reports, inventories) ◆ National and state databases (government budgets, patents, census data) ◆ Survey results

Method	Typical Questions about Office of Science (SC) Programs	Distinctive Data, Units of Measure, Analytic Tools
Key Indicators (Continued)	<ul style="list-style-type: none"> ♦ How has the profile of scientists at DOE laboratories changed over time? ♦ What evidence is there that long-run target goals will be met? 	Units of Measure <ul style="list-style-type: none"> ♦ Dollars of costs ♦ Counts of outputs ♦ Outputs per unit cost Analytic Tools <ul style="list-style-type: none"> ♦ Statistical analysis software

Table 7. Matching Methods and Study Objectives

Primary Study Objectives	Frequently Used Methods (Key at Bottom)												
To Gain Management Insights													
♦ Assess quality of research	2												
♦ Identify pathways of knowledge flows	5, 7												
♦ Identify trends and emerging areas in research	5, 6												
♦ Assess S&T program performance	2, 8, 9, 11												
♦ To understand relationships among inputs, program mechanisms, and outputs	1, 3, 10, 11												
To Understand Issues/Challenges, e.g.,													
♦ Are collaborative relationships working?	7, 8, 10												
♦ What influences success?	8, 10												
♦ Who is using research results? For what?	1, 4, 5, 7, 8												
♦ What is the relationship among research areas?	5, 6												
To Assess Core Competencies & Unique Contributions, e.g.,													
♦ To what extent have past investments paid off?	4, 8, 9												
♦ Are users satisfied with special facilities?	8												
♦ What are the spheres of influence of DOE researchers?	7												
To Communicate/Explain Research, e.g.,													
♦ Interest/educate/inform a larger audience	1, 4												
♦ Generate program metrics (GPRA)	2, 5, 8, 9, 11												
KEY: <table> <tr> <td>(1) Classic, Interview- and Records-Based Method</td><td>(7) Sociometric/Social Network Methods</td></tr> <tr> <td>(2) Expert Review Method</td><td>(8) User/Participant Survey Method</td></tr> <tr> <td>(3) Process-Outcomes Method</td><td>(9) Benefit-Cost (Cost-Benefit) Methods</td></tr> <tr> <td>(4) Historical Tracing Method</td><td>(10) Statistical/Econometric Methods</td></tr> <tr> <td>(5) Bibliometrics and Citation Methods</td><td>(11) Key Indicator Method</td></tr> <tr> <td>(6) Content-Analysis Methods (Including Co-Word Method, Data Tomography, Textual Data Mining, and Visualization)</td><td></td></tr> </table>		(1) Classic, Interview- and Records-Based Method	(7) Sociometric/Social Network Methods	(2) Expert Review Method	(8) User/Participant Survey Method	(3) Process-Outcomes Method	(9) Benefit-Cost (Cost-Benefit) Methods	(4) Historical Tracing Method	(10) Statistical/Econometric Methods	(5) Bibliometrics and Citation Methods	(11) Key Indicator Method	(6) Content-Analysis Methods (Including Co-Word Method, Data Tomography, Textual Data Mining, and Visualization)	
(1) Classic, Interview- and Records-Based Method	(7) Sociometric/Social Network Methods												
(2) Expert Review Method	(8) User/Participant Survey Method												
(3) Process-Outcomes Method	(9) Benefit-Cost (Cost-Benefit) Methods												
(4) Historical Tracing Method	(10) Statistical/Econometric Methods												
(5) Bibliometrics and Citation Methods	(11) Key Indicator Method												
(6) Content-Analysis Methods (Including Co-Word Method, Data Tomography, Textual Data Mining, and Visualization)													

3.10 Develop a Preliminary Study Design

Based on the decisions that have been made about the purpose and scope of the case study and the methods that might be used, the next step is to develop a preliminary study design. The assistance of a qualified analyst may be helpful for those less familiar with social science research methods.

The research design is the logic that links data collection and conclusions to the initial questions of the study. The research design is a blueprint of what questions to study, what data are relevant, what data to collect, and how to analyze the results (Yin 1984:18-20). The preliminary study design should specify (Yin 1984):

- ◆ The study's questions
- ◆ Its propositions, if any
- ◆ The unit of study
- ◆ The time frame
- ◆ The logic linking the data to the propositions
- ◆ The principal methods of data collection and analysis to be used
- ◆ The criteria for interpreting the findings.

The write-up of the preliminary design should address each of the factors considered in planning the study:

- ◆ Purpose and scope
- ◆ Audience
- ◆ Questions to be answered
- ◆ Time frame
- ◆ Case selection
- ◆ Collaborators and partnerships
- ◆ Methods
- ◆ Schedule and resources.

This preliminary design should be prepared in a form that can be reviewed by the Office of Science, collaborating organizations, and an external review team.

3.11 Conduct an Initial Internal and External Review

Internal and external reviews following the development of a preliminary study design provide a variety of benefits to the sponsors and managers of case studies. Internal reviews capitalize on the special knowledge of staff about the organization, its context, and the particular case being studied. They provide an opportunity to make sure the proposed case study is appropriately framed, focused on important questions, and designed to access the best sources of information. Internal reviews also have the beneficial effect of raising awareness of the study, and can lay the groundwork for receptivity to study findings.

Depending upon the formality of the internal review, it may be necessary to document the reviewers' comments and how they have been addressed, including circulation of a revised

proposal. It is important to take advantage of the reviewers' comments in revising the proposed study design and to let the reviewers know that their input has been given careful consideration.

A good review panel includes people familiar with the:

- ◆ Project or program being studied
- ◆ Methods of data collection and analysis
- ◆ Organizational themes and contributing elements
- ◆ Broad context in which the values and impacts can be assessed
- ◆ Other case studies that have been conducted.

The preliminary design may also require external review. The well-established Office of Science process for external reviews serves as a model for this process, with modifications as necessary. To the extent possible, it is helpful for the reviewers to be familiar with the problem being addressed and knowledgeable about the topic and proposed methods.

3.12 Commission the Analysts and Finalize the Study Design


Finalizing the approach and developing the detailed study design is best undertaken as a collaborative effort between the program manager and the analyst(s) who will conduct the research and prepare the research report, in consultation with collaborators and partners. The preceding elements of the planning phase should have provided a solid basis for determining the nature and extent of assistance needed. The selection and use of analysts is a project management decision that will vary depending on the conditions of the case study, the particular method(s) and approach being used, and the guidance provided by the study sponsor and the Office of Science.

Once the analysts have been selected, they should work with the program manager to develop a detailed study design based on the needs delineated in the preliminary design. The detailed study design should include:


- ◆ Elaboration of each of the items in the preliminary design
- ◆ A statement of the key research questions and the conceptual or theoretical framework or model upon which they are based
- ◆ A formal plan for collecting the data that provides
 - Identification of the cases to be studied
 - Specification of the data to be collected and the sources of that data
 - Procedures to address issues of consent, confidentiality, privacy, and protection of proprietary information
 - Protocols for contacting those involved in or representing the case
 - Protocols for collecting the data
 - Protocols for verifying data quality
 - “Table shells” for arranging the data and a chart that maps data to questions
- ◆ Specification of the data analysis tools and methods to be used
- ◆ Identification and a general description of the deliverables (reports, presentations, and data sets)

- ◆ Protocols for interaction with collaborators, including review and release of information about the study and its findings
- ◆ Personnel, their qualifications, and their roles and responsibilities
- ◆ A schedule
- ◆ A budget by major study component.

One of the greatest challenges in case study research is the creation of quantitative measures that meet the desired standards of good research practice and rigor, yet are feasible in terms of data collection. This is particularly important if the case study is attempting to develop time series or cross-case comparisons, which depend upon consistency and completeness for validity. A special effort to identify the data needed for counts, scales, and categorizations in the design phase can save much frustration and backtracking during study implementation. Experienced researchers recommend the development of templates, “table shells,” and charts that make explicit what data will be obtained, in what form, and from which sources. Larsson (1993), Thorne (1994), Swanson and Chapman (1994), and Bozeman et al. (1998) provide useful discussions of this process.

 **Minimize backtracking** by identifying the specific data needed to construct desired quantitative measures.

Once the detailed research design has been prepared, it may need to be reviewed by a Human Subjects Committee (if it involves human subjects in the research) or by the Office of Management and Budget (if it involves a survey) before fieldwork can begin. The program manager and the analyst must agree about access to primary data and information about data sources. Depending upon the nature of the study and the need for independent validation of findings, confidentiality agreements for third party access to the primary data may also be needed. Serious legal and ethical issues may be involved, so it is important to take this aspect of case study research seriously and to obtain informed advice about requirements and good practice. Problems can be avoided with proper attention to these sensitive issues.

 **Take ethical and privacy issues seriously.** Seek out and follow informed advice.


This page intentionally left blank

4. MANAGING CASE STUDY IMPLEMENTATION

Managing a case study is similar to managing any research project, but there are some important differences. Office of Science case studies deal with policy- and performance-related issues. They frequently involve collecting data from people, and almost always require notification of and interaction with the people and organizations associated with the case(s) being studied. Ensuring that these interactions proceed smoothly and in accordance with legal and ethical requirements is the joint responsibility of the program manager and the analyst. In addition to ensuring that the case study proceeds on schedule, within budget, and in accordance with good research practices, the program manager may be called upon to facilitate interactions, respond to questions or issues about the case study, and provide status reports to study sponsors and key stakeholders.

4.1 Provide Management Updates

Because they address policy- and performance-related issues, Office of Science case studies are often relatively visible, both internally and externally. Routine progress reports and briefings to key stakeholders and managers should call out issues of particular interest or sensitivity, and provide updates on emerging findings. Maintaining an up-to-date presentation that describes the key aspects of the study and provides a status update will make it easier to keep interested parties informed. Regular briefings to study sponsors and interested stakeholders can head off problems and generate support for the study and its findings.

 **Head off issues and build support by keeping upper management and stakeholders informed through regular briefings.**

4.2 Monitor Common Data Collection and Analytic Difficulties

Ethical and legal issues should be identified and addressed in the detailed study design. They must be monitored during study implementation. Compliance with disclosure and consent procedures must be documented. Although qualified analysts should be familiar with these requirements, program managers need to make sure procedures are followed and appropriate records maintained.


Because Office of Science case studies typically draw upon information held by others, problems with focus and the availability, accessibility, and quality of data generally pose the greatest challenge in conducting high quality case studies, as illustrated in Table 8. Monitoring progress on data collection, documentation, and analysis can reinforce good research practices and identify problems in time to implement corrective measures. Reviewing interim results can reveal unanticipated findings that may indicate design or data collection problems or an important discovery that warrants a change in course or strategy. Yin (1984) provides an excellent discussion of the issues associated with case study research implementation. The Research Value Mapping project at Georgia Institute of Technology has also produced a number of papers discussing these issues (Bozeman et al. 1998).

Table 8. Aspects of Data Collection Typically Causing the Greatest Difficulty

<p>Developing and maintaining a clear framework and focus</p> <ul style="list-style-type: none"> ◆ Matching scope and budget so that an adequate density of data is assembled on key topics ◆ Specifying and then focusing on collecting the data most pertinent to the study's purpose and audience <p>Addressing privacy and confidentiality concerns and fulfilling confidentiality commitments</p> <ul style="list-style-type: none"> ◆ Establishing and adhering to appropriate informed consent procedures ◆ Maximizing the research value of data while protecting respondents' privacy and interests <p>Contacting the people who need to be interviewed</p> <ul style="list-style-type: none"> ◆ Developing an appropriate sample frame (the list of all people who might be interviewed), particularly for survey-based methods ◆ Obtaining OMB clearance for survey questionnaires ◆ Finding and notifying the people who have been selected for interviews ◆ Scheduling appointments with those who are to be interviewed <p>Obtaining the secondary data needed for the analysis</p> <ul style="list-style-type: none"> ◆ Finding the pertinent data archives or sources ◆ Obtaining permission to access or copy the data ◆ Obtaining the data in a format that is compatible with the analytic tools ◆ Dealing with poor quality data (missing or incomplete data, sloppy records) <p>Maintaining consistency in data collection across multiple cases, field workers, or time periods</p> <ul style="list-style-type: none"> ◆ Balancing modifications to capitalize on new insights against the additional time and effort required to collect the new information for all cases/ time periods ◆ Validating codes and coding procedures with inter-rater reliability checks <p>Documenting and verifying the source and quality of the data</p>
--


4.3 Maintain Coordination among Collaborators and Partners

It is important to maintain the collaborative approach established during the planning phase by engaging partners in an ongoing review of progress and findings and by following agreed-upon protocols. Key factors in successful collaborations are timely and open communication, and sensitivity to collaborators' concerns, constraints, and needs – particularly with regard to review and release of case study findings.

 Engage collaborators early and honor agreements about the release of case study findings.

4.4 Track Key Issues in Case Study Data Analysis

A common problem in case studies is failing to allocate adequate time and budget for data analysis and report preparation. Program managers who are unfamiliar with the data collection and analytic techniques being used in the study might want to obtain expert assistance in overseeing the analysis and report preparation phase of the study.

 **Don't short-change analysis and report preparation by spending too much time and money collecting data.**

The large volume of qualitative and quantitative data typical of case studies poses challenges for analysis and report preparation. These challenges include:

- ◆ *Maintaining objectivity.* This is a significant challenge in case study research. Program managers need to be alert to this problem and reinforce the importance of objectivity in both data analysis and report preparation. They should challenge the analysts to acknowledge potential biasing factors and to demonstrate what steps they are taking to ensure objectivity, for example by incorporating sufficiently robust counterfactuals.
- ◆ *Keeping focused.* An important benefit of the case study approach is that it is open to exploration and discovery. However, if not carefully managed, case studies can become so broad that the density of data is insufficient to support convincing conclusions. Program managers can help maintain an appropriate focus by being realistic about what can be addressed with the available time and money, and by having the analyst map out how the data will support descriptions and conclusions.
- ◆ *Collecting data that have internal and external validity.* It is often difficult to find data that accurately and comprehensively represent the measures of greatest interest in case studies of science (for example “excellence” in science). A combination of creativity and rigor is needed to develop data that best achieve this goal.
- ◆ *Coding, categorizing, and analyzing qualitative data* in consistent and theoretically valid ways. This involves:
 - Avoiding unwarranted precision
 - Using parametric and non-parametric statistics appropriately
 - Specifying and documenting data sources and analytic procedures.
 - Applying analytic procedures consistently and systematically
 - Validating objectivity and consistency of coding and interpretation
 - Acknowledging the limitations of the data and analysis.
- ◆ *Addressing alternative perspectives and interpretations* by:
 - Providing controls
 - Developing counterfactuals
 - Acknowledging and addressing alternative explanations
 - Conducting “future without” analyses.
- ◆ *Confirming accuracy in fact and interpretation* with the subjects of the study. It is common practice in qualitative data analysis to ask those who provided the information and/or who are knowledgeable about the events being studied to review the draft report and validate facts and interpretations. These reviews add considerable credibility and are often crucial to gaining acceptance of results.

4.5 Oversee Preparation of the Case Study Report and Oral Presentation Materials

Case study reports and oral presentations should reflect audience needs. While protecting the independence of the analyst with regard to study findings and results, the program manager should provide guidance about the expected format and style of case study reports and presentations, following consultation with any collaborators on the study. *The Craft of Research* (Booth et al. 1995) provides clear advice on writing effective research reports.

The process begins with a review of the information assembled to design the case study. The program manager and analyst should clarify (based on Booth et al. 1995):


- ◆ The principal, secondary, and tertiary audiences for the report
 - Who they are (policy makers, managers, professions in the field, science attentive public, general public)
 - What they already know about the topic
 - How they think about the topic/problem
 - Their key questions or concerns
 - In what forum they will encounter the report
 - What format or mode of presentation they expect (if any)
- ◆ The purpose and scope of the study
- ◆ Key positioning issues (linked to organizational themes and collaborations)
- ◆ The key questions driving the research
- ◆ How and why the cases were chosen.

Reports, articles, and presentations should feature the case study results. The methodology should be described clearly, but should not be the primary focus of the final product. The technical report should include:


- ◆ An executive summary that communicates the bottom line to the primary audience
- ◆ A description of the research methods that a qualified researcher could follow to replicate the study
- ◆ Clear, concise explanations of key assumptions
- ◆ Strong, crisp text that highlights key points
- ◆ Evidence to support conclusions and refute alternative explanations, presented in a way that anticipates audience preconceptions, doubts, and confusion
- ◆ Liberal use of graphical and visual representations⁷ that reinforce the message and add interest by
 - Summarizing multiple attributes and comparisons
 - Accurately representing the data and key interpretations
- ◆ Complete bibliographic information for secondary sources and reference material
- ◆ Appendices presenting key data tables.

⁷ See <http://www.math.yorku.ca/SCS/Gallery/noframes.html> for a good discussion and ideas on graphics.

If there are multiple audiences, write derivative reports or articles to reach each targeted group. Analysts skilled at communicating with other analysts may not communicate well with other audiences. Skilled technical writers can tailor case study findings to target groups and help tell the story effectively. The specifications of derivative reports will depend upon the nature of the study, the intended target audiences, and the needs of any collaborators or partners in the study.

 Use a technical writer to help convey case study findings effectively.

In addition to the written reports, materials to support oral and visual presentations should be considered an integral part of the study deliverables. Presentation requirements generally range from a single visual that captures the gist of the study to a detailed presentation suitable for delivery in approximately 30 to 45 minutes. The supporting presentation materials should reflect, and be keyed to, key points in the technical and derivative reports. Any valid deviations should be noted and clearly explained.

 Require analysts to provide materials that support oral and visual presentations.

4.6 Coordinate the Review, Revision, and Finalization of Study Report(s)

Review of the study's results provides an opportunity to identify and correct errors of fact or interpretation and to enhance the credibility of the final report. Internal review provides an additional opportunity to capitalize on the insights of staff with first-hand knowledge of the case as well as the context within which it occurred. It also builds awareness of the study and its findings.

It is courteous to provide the subjects of the case study the first opportunity to review the report. It is especially important to provide collaborators and study participants who may be affected by the release of the information an opportunity to review sensitive findings prior to their release in a public setting, giving them an opportunity to understand the findings and prepare response statements.

Depending upon the visibility of the case study, a formal external review may also be conducted following the Office of Science's procedures. Both internal and external reviewers should be given clear guidance about how and when to submit their comments and a clear explanation of how their comments will be used and addressed.

Once the reviews are completed, the program manager, analyst, and collaborators should decide how to address the comments and revise the report(s) and oral presentation materials. The power of the case study to educate and enhance policy and management can be diminished if the findings are presented without adequate review. If a thorough and even-handed review has been conducted, the case study results are more likely to be used and applied in future decisions. The usual Office of Science procedures should be followed in publishing and distributing the final report(s). Unless there are particular reasons to restrict distribution of case study reports, they should also be posted on the Office of Science website.

The data collected during the course of the study, along with analytic results, should be archived and/or stored consistent with the protocols established in the detailed study design. Particular care should be taken to ensure that the privacy, confidentiality, and proprietary information of study participants will continue to be adequately protected even after the study is completed.

This page intentionally left blank

5. COMMUNICATING AND APPLYING CASE STUDY RESULTS

Ultimately, the value of the case study is realized through clear communication of the findings and lessons learned. All too often, the final steps of communication and application are taken for granted. Researchers may assume that “if you build it, they will come,” and that a completed report will automatically accomplish its intended goal. Sometimes this is the case, but usually additional planning and effort is required.


The case study process culminates with communication of the results to intended audiences and application of the findings to their intended purposes. These include:

- ◆ Supporting programmatic decision-making
- ◆ Communicating impacts and accomplishments to interested stakeholders
- ◆ Responding to special requests for information
- ◆ Informing policy
- ◆ Meeting specific legislative requirements for performance information

5.1 Gauge the Target and Plan the End-Game


Approach the end game – communication and application – with the same rigor and planning as the research and analysis itself:

- ◆ Think creatively. A variety of materials can augment the final report: Simple to complex viewgraphs, real-time computer-aided sessions, summary issue papers, and possibly white papers addressing particular subsets of issues discovered through the case study process.
- ◆ Don't limit yourself to reports. Brochures, conference presentations, workshops, or published technical or media articles may be more effective with a general audience.
- ◆ Involve team members. Depending upon the venue and audience, analysts, collaborators, and partners can add credibility and visibility by participating in presentations.
- ◆ Think cyberspace. The internet, including the Office of Science website, provides a powerful tool for reaching many potential stakeholders, in part because it makes research visible and accessible for a longer period of time. Postings on the internet can be “hot-linked” to other resource material, and can make bibliographies, abstracts, summaries, and other supporting information available inexpensively. Many more possibilities exist. The purpose of the case study, and the interests and needs of the target audiences, must be central to the preparation of a plan for the end game.

 Approach the end-game with the same rigor that you approach the analysis!

Your end game needs to reflect the different interests and levels of technical competency among your audiences. This is particularly important when the case study has a significant technical component. The following illustrates the potential interests of a case study's audiences:


- ◆ *Scientists and Researchers:* Trends and new directions in

 Keep your eye on the ball! Know your intended applications and the target audience(s).

- fields of science, new discoveries, technical communication pathways, funding/sponsorship trends, peer progress and capabilities, facility plans
- ◆ *Project Managers:* Process improvements, planning and scheduling, project interrelationships, resource allocations or reallocations, technical tradeoffs, efficiency issues, stewardship issues, response to new discoveries, benchmark information
 - ◆ *Legislative and Office of Management and Budget staff:* Overall value and impacts, alignment with national policies and shifting emphases, portfolio balance, budget sensitivities and uncertainties, interplay across institutions, performance relative to commitments, linkages between basic and applied research communities, geographic distribution of research
 - ◆ *The Science-Attentive Public and Various Stakeholders:* Insights into the process of science, scientific discoveries, possible areas of future discoveries, future scenarios, and possible applications of the research
 - ◆ *Senior Office of Science and DOE management:* Many of the same items listed for the legislative and Office of Management and Budget staff, plus general stewardship issues, field management and site issues, and management tools and efficiency issues.

5.2 Strengthen the Message: Build On Case Study Findings

Individual case studies provide valuable evaluative and research information that can inform managers and illuminate policy. Collections of case studies provide even more value. As with any research, a larger sample is generally better than a smaller one. Single case studies seldom provide definitive answers to the types of questions that concern the Office of Science and its stakeholders. Findings gain credibility if they are replicated in multiple studies. New insights can be gained by comparing results across cases, as is cogently discussed by Yin (1984). In addition, the design of new studies can be informed by review of those conducted previously. The ability to gain this advantage depends upon having an accessible, searchable archive of case study reports and data, something that is increasingly feasible through internet posting of case study reports.

 Themes reinforced through multiple case studies offer considerably stronger messages..... the whole is greater than the sum of its parts.

Organizational themes, such as those discussed in Chapter 2 and illustrated in Table 2, provide a basis for linking multiple case studies together and for identifying the broad, organizationally significant areas where the findings of the case study may apply. It is not difficult to see that it is easier to apply a particular case study's findings when they reinforce common points, in this case themes. Indeed, care must be taken not to allow such correspondence to override appropriate caution in interpreting the strength of study findings.


5.3 Share Findings with Secondary and Tertiary Audiences

A final consideration is sharing the case study findings and methodology issues with secondary and tertiary audiences. The initial planning for the study will have addressed the basic issues of audiences, stakeholders, collaborators, and partners. Each of these groups may have different information needs and interests in the study results. The program manager and analyst, along with Office of Science communications specialists, can use information and documentation collected throughout the case study to prepare base materials that can be modified to meet these

different needs. Each presentation should be tailored for the particular audience. It is useful to prepare, review, and have on hand:

- ◆ An abstract
- ◆ An executive summary
- ◆ Press release(s)
- ◆ Written reports (long and short versions)
- ◆ Presentation slides or CDs (targeting various audiences)
- ◆ Database with documentation.

Listserves of related organizations are a simple way of communicating the availability of new reports and presentation materials, or if the reports are posted on the internet, to provide access to the case study materials themselves. Presentation of results and distribution of brochures and other materials at conferences, workshops, and meetings provide valuable notification, too.

 The Internet can be an effective way to communicate with a broad audience, but good materials are needed to do this well.

Although communication with these audiences is typically a secondary objective of research, these extra steps can strengthen future case study efforts by contributing to the broader body of knowledge in both the technical area of the study and in case study methodology. Ultimately, the value of the study results from clear communication of the findings and the lessons learned from conducting the research.

This page intentionally left blank

6. REFERENCES

- Austin, David, and Molly Macaulay. 2000. Estimating Future Consumer Benefits from ATP-Funded Innovation: The Case of Digital Data Storage. NIST GCR 00-790. Washington, DC: Resources for the Future. (Available on-line at http://www.atp.nist.gov/eao/eao_pubs.htm).
- Balch, P. and S. Brunak. 1998. *Bioinformatics: The Machine Learning Approach*. Cambridge, MA: MIT Press.
- Barzun, Jacques, and Henry F. Graff. 1992. *The Modern Researcher*. 5th Edition. New York: Harcourt Brace Jovanovich.
- Blalock, H. M., Jr. 1972. *Social Statistics*. Second edition. New York: McGraw-Hill.
- Blau, J. 1978. Sociometric Structure of a Scientific Discipline. In *Research in the Sociology of Knowledge, Sciences, and Art: An Annual Compilation of Research I*. Edited by R. Jones. Greenwich, CT: JAI Press.
- Boardman, Anthony E. (ed). 1996. *Cost Benefit Analysis: Concepts and Practice*, 1st Edition, Englewood Cliffs, NJ: Prentice-Hall Inc.
- Booth, Wayne C., Gregory G. Colomb, and Joseph M. Williams. 1995. *The Craft of Research*. Chicago: The University of Chicago Press.
- Borgman, Christine L., et al. (eds). 1990. *Scholarly Communication and Bibliometrics*. Newbury Park, CA: Sage Publications, Inc.
- Bozeman, B. 1993. Peer Review and Evaluation of R&D Impacts. In *Evaluating R&D Impacts: Methods and Practice*. Edited by B. Bozeman and J. Melkers. Norwell, MA: Kluwer Academic Publishers.
- Bozeman, B., and F. Donez. 1996. Brookhaven National Laboratory, Superconducting Materials and Magnet Technology. Prepared for Sandia National Laboratory. Atlanta, GA: Georgia Institute of Technology.
- Bozeman, B, J. Rogers, D. Roessner, H. Klein, and J. Park. 1998. *The R&D Value Mapping Project: Final Report*. Prepared for the U.S. Department of Energy, Office of Basic Energy Sciences. Atlanta, GA: Georgia Institute of Technology.
- Breitzman, A., and F. Narin. 1996. Linkage between Patents and DOE BES Supported Papers. CHI Contract No. 9521.
- Callon, M., J. Courtial, and F. Laville. 1991. Co-Word Analysis as a Tool for Describing the Network of Interactions Between Basic and Technological Research: The Case of Polymer Chemistry. *Scientometrics* 22(1):155-205.
- Cohen, Wesley M. 1995. Empirical Studies of Innovative Activity. In *Handbook of the Economics of Innovation and Technological Change*. Edited by Paul Stoneman. Cambridge: Blackwell.

- The Committee on Science, Engineering, and Public Policy (COSEPUP). 1999. *Evaluating Federal Research Programs: Research and the Government Performance and Results Act*. Produced with support from National Academy of Sciences, National Academy of Engineering, Institute of Medicine. Washington, DC: National Academy Press. (Can be ordered on-line from www.nap.edu/catalog/6416.html)
- Cozzens, S. 1989. Literature-Based Data in Research Evaluation: A Manager's Guide to Bibliometrics. Report to the National Science Foundation.
- David, Paul A., David Mowery, and W. Edward Steinmueller. 1992. Analysing the Economic Payoffs of Basic Research. *Economics of Innovation and New Technology* 2(1):73-90.
- Dillman, Don A. 2000. *Mail and Internet Surveys: The Tailored Design Method*. Second Edition. New York: John Wiley and Company.
- Dillman, Don A. 1978. *Mail and Telephone Surveys*. New York: John Wiley and Sons.
- Ehlen, Mark A. 1999. Economic Impacts of Flow-Control Machining Technologies: Early Applications in the Automotive Industry. NISTIR 6373. Gaithersburg, MD: NIST. (Available on-line at http://www.atp.nist.gov/eao/eao_pubs.htm).
- Eisenhardt, K.M. 1989. Building Theories from Case Study Research. *Academy of Management Review* 14(4):532-550.
- Feldman, Maryann, and Maryellen Kelley. 2001. The Case for Government R&D Additionality: An Investigation of the Advanced Technology Program Selection Process. In *The Advanced Technology Program: Assessing Outcomes*. National Research Council/National Academies of Science. Washington, DC: National Academy Press.
- Feldman, Maryann, and Maryellen Kelley. March 2001. Winning an Award from the Advanced Technology Program: Pursuing R&D Strategies in the Public Interest and Benefiting from a Halo Effect. National Institute of Standards and Technology, NISTIR 6577. (Available on-line at <http://www.atp.nist.gov/eao/ir-6577.pdf>)
- Fiedler, J. 1978. *Field Research: A Manual for Logistics and Management of Scientific Studies in Natural Settings*. San Francisco: Jossey-Bass.
- Freeman, Linton C. (undated). *Visualizing Social Networks*. (Available on-line at <http://tarski.ss.uci.edu/vis.html>).
- Garton, Laura, Caroline Haythornthwaite, and Barry Wellman. 1997. Studying Online Social Networks. *Journal of Computer-Mediated Communication* 3 (1). (Available on-line at <http://www.ascusc.org/jcmc/vol3/issue1/garton.html>).
- Geisler, E. 2000. The Metric of Peer Review. In *The Metrics of Science and Technology*. Edited by E. Geisler. Westport, CT: Quorum Books.
- Geisler, E. 1995. An Integrated Cost-Performance Model of Research and Development Evaluation. *Omega* 23(3):Appendix A.
- Geisler, E. (ed.) 2000. *The Metrics of Science and Technology*. Westport, CT: Quorum Books. (Can be ordered from the Greenwood Publishing Group (www.greenwood.com)).
- Georghiou, G., W. Giusti, H. Cameron, and M. Gibbons. 1988. The Use of Co-nomination

- Analysis in the Evaluation of Collaborative Research. In *Handbook of Quantitative Studies of Science and Technology*. Edited by A.F.J. Van Raan. North Holland: Elsevier Science.
- Gleick, James. 1987. *Chaos: Making a New Science*. New York: Viking Press.
- Gramlich, Edward A. 1997. *A Guide to Benefit-Cost Analysis*. Second edition. Prospect Heights, IL: Waveland Press.
- Gregory, Jane, and Steve Miller. 1998. *Science in Public: Communication, Culture, and Credibility*. Cambridge, MA: Perseus Publishing.
- Grol, R., and M. Lawrence (eds). 1995. *Quality Improvement by Peer Review*. New York: Oxford University Press.
- Harrison, Michael I. 1987. *Diagnosing Organizations: Methods, Models, and Processes*. Applied Social Research Methods Series. Volume 8. Beverly Hills, CA: Sage Publications.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- Kish, Leslie. 1965. *Survey Sampling*. New York: John Wiley and Sons.
- Klein, H., and B. Bozeman. 1998. Los Alamos National Laboratory, Thermoacoustic Heat Engines. Atlanta, GA: Georgia Institute of Technology.
- Kostoff, R. 1997. The Principles and Practices of Peer Review. *Science and Engineering Ethics* (Special Issue on Peer Review) 3:19-34.
- Kostoff, Ronald N. 1997. *The Handbook of Research Impact Assessment: Seventh Edition*. Defense Technical Information Center, Report No. ADA296021. Office of Naval Research. (Available on-line at <http://www.dtic.mil/dtic/kostoff/handweb7index.htm>).
- Krishnaiah, P., and C. Rao. 1988. *Handbook of Statistics 6: Sampling*. North-Holland: Elsevier Science Publisher. (Abstract and ordering information for series available on-line at <http://www.elsevier.nl/inca/publications/store/5/2/4/3/8/9/>).
- Kroll, P., G. Ault, and F. Narin. 1998. Tracing the Influence of Basic Scientific Research on Biotechnology Patents: A Case Study. *Patent World* March:38-46.
- Larsson, R. 1993. Case Survey Methodology: Quantitative Analysis of Patterns Across Case Studies. *Academy of Management Journal* 36(6):1515-1546.
- Link, A. 1997. Economic Evaluation of Radiopharmaceutical Research at NIST. 97-2 Planning Report. (Available on-line at www.nist.gov/director/prog-ofc/report97-2.pdf)
- Lohr, S.L. 1999. *Sampling: Design and Analysis*. Pacific Grove, CA: Duxbury Press.
- McMillan, G., F. Narin, and D. Deeds. 2000. An Analysis of the Critical Role of Public Science in Innovation: The Case of Biotechnology. *Research Policy* 29(1, January):1-8.
- Mansfield, E. 1991. Academic Research and Industrial Innovation. *Research Policy* 20(1):1-12.
- Martin, Sheila A., Daniel L. Winfield, Anne E. Kenyon, John R. Farris, Mohan V. Bala, and

- Taylor H. Bingham. 1998. *A Framework for Estimating the National Economic Benefits of ATP Funding of Medical Technologies*. NIST GCR 97-737. Research Triangle Park, NC: Research Triangle Institute, Center for Economic Research. (Available on-line at <http://www.atp.nist.gov/eao/97-737/contents.htm>).
- Maxwell, Robert A., and Shohreh B. Eckhardt. 1990. *Drug Discovery: A Case Book and Analysis*. Clifton, NJ: Humana Press.
- Miller, R. 1992. The Influence of Primary Task on Research-and-Development Laboratory Evaluation - A Comparative Bibliometric Analysis. *R&D Management* 22(1):3-20.
- Miller, R. and A. Manseau. 1996. Bibliometric Indicators and the Competitive Environment of Research-and-Development Laboratories. *Scientometrics* 36(3):421-433.
- Morgan, David L. 1997. *Focus Groups As Qualitative Research*. Second edition. Qualitative Research Methods Volume 16. Newbury Park, CA: Sage Publications.
- Morse, Janice M. (ed). 1994. *Critical Issues in Qualitative Research Methods*. London: Sage Publications.
- Narin, F. 1994. Patent Bibliometrics. *Scientometrics* 30(1):147-155.
- Narin, F. 1997. The Linkage between Patents and the Scientific Literature. In *Knowledge Production, Patents and Technological Intelligence*. Proceedings of a Conference on the Assessment and Use of Scientific and Technological Knowledge. Canberra, Australia.
- Narin, F., K. Hamilton, and D. Olivastro. 1995. Linkage between Agency-Supported Research and Patented Industrial Technology. *Research Evaluation* 5(3 December):183-187.
- National Research Council (NRC), Committee on the Department of Energy-Office of Science and Technology's Peer Review Program. 1999. *Peer Review in Environmental Technology Development Programs*. National Academy Press. (Available on-line at www.nap.edu/catalog/6408.html).
- National Academies of Sciences. *Beyond DiscoveryTM: The Path from Research to Human Benefit*. (Available on-line at <http://www4.nas.edu/beyond/beyonddiscovery.nsf>. Last updated on 1/31/2001.).
- National Center for Case Study Teaching in Science: Case Study Collection. (Available on-line at <http://ublib.buffalo.edu/libraries/projects/cases/ubcase.htm>. Last updated on 03/08/01.)
- National Science Board. 2000. *Science & Engineering Indicators - 2000*. Washington, DC: U.S. Printing Office (Available on-line at www.nsf.gov/sbe/srs/seind00/start.htm).
- Newman, M.E. J. 2001. The Structure of Scientific Collaboration Networks. *Proceedings of the National Academy of Science* 98(2):404-409.
- OECD. 2001. *Main Science and Technology Indicators 2000*. (Available on-line at www.sourceoecd.org/bookshop).
- Paradis, James G. and Muriel L. Zimmerman. 1997. *The MIT Guide to Science and Engineering Communication*. Cambridge: MIT Press.
- Patton, M. 1990. *Qualitative Evaluation and Research Methods*. Newbury Park, CA: Sage

Publications.

- Perko, J., and F. Narin. 1997. The Transfer of Public Science to Patented Technology: A Case Study in Agricultural Science. *The Journal of Technology Transfer* 22 (3):65-72.
- Policy Analysis Applications of REMI Economic Forecasting and Simulation Models. 1995. *International Journal of Public Administration* 18(1):13-42.
- Popper, S. 1995. *Economic Approaches to Measuring the Performance and Benefits of Fundamental Science*. RAND PM-409-OSTP. July.
- Powell, J. 1999. *Business Planning and Progress of Small Firms Engaged in Technology Development through the Advanced Technology Program*. NIST 6375. October. (Available on-line at www.atp.nist.gov/eao/ir-6375.pdf).
- Ragin, Charles C., and Howard S. Becker (eds). 1992. *What is a Case? Exploring the Foundations of Social Inquiry*. Cambridge: Cambridge University Press.
- Research Value Mapping (RVM) Program. Institute for Policy Research and Development School of Public Policy, Georgia Institute of Technology. Barry Bozeman, RVM Program Director. (See <http://rvm.pp.gatech.edu>).
- Rhodes, Richard. 1986. *The Making of the Atomic Bomb*. New York: Simon and Schuster.
- Risch, J.S., D.B. Rex, S.T. Dowson, T.B. Walters, R.A. May, and B.D. Moon. 1999. The STARLIGHT Information Visualization System. *Readings in Information Visualization*. Edited by S. Card, J. Mackinlay, and B. Shneiderman. Pp. 551-560. San Francisco: Morgan Kaufmann.
- Roessner, David. 2000. Quantitative and Qualitative Methods and Measures in the Evaluation of Research. *Research Evaluation* 8(2):125-132.
- Rogers, Juan D. Undated. University of California-Lawrence Berkeley Laboratory Nuclear Magnetic Resonance Spectroscopy; A Case Study of the R&D Value Mapping Project, Case 97-07. Sponsored by the Department of Energy, Office of Basic Energy Sciences. Georgia Institute of Technology. (Available on-line at <http://rvm.pp.gatech.edu/cases/pines-case.html>).
- Ruegg, R. 2001. Taking a Step Back: An Early Results Overview of Fifty ATP Awards. In *The Advanced Technology Program: Assessing Outcomes*. National Research Council/National Academies of Science. Washington, DC: National Academy Press.
- Ruegg, R. 1997. Economic Methods. In *CRC Handbook of Energy Efficiency*. Edited by F. Kreith and R. West. Boca Raton: CRC Press.
- Ruegg, R., and H. Marshall. 1990. *Building Economics: Theory and Practice*. New York: Van Nostrand Reinhold.
- Science Policy Research Unit (SPRU). 1972. *Success and Failure in Industrial Innovation*. London: Center for the Study of Industrial Innovation.
- Scott, John. 1991. *Social Network Analysis*. London: Sage Publications.
- Sieber, Joan E. (ed.). 1982. *The Ethics of Social Research: Surveys and Experiments*. New

York: Springer.

SPIRE™ website: <http://www.pnl.gov/infoviz> .

Sudman, S., and N.M. Bradburn. 1982. *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco, CA: Jossey-Bass.

Swanson, Janice M., and Linda Chapman. 1994. Inside the Black Box: Theoretical and Methodological Issues in Conducting Evaluation Research Using a Qualitative Approach. In *Critical Issues in Qualitative Research Methods*. Edited by Janice M. Morse. Pp. 66-93. London: Sage Publications.

Thorne, Sally. 1994. Secondary Analysis in Qualitative Research: Issues and Implications. In *Critical Issues in Qualitative Research Methods*. Edited by Janice M. Morse. Pp. 263-279. London: Sage Publications.

Toole, Andrew A. 2000. The Impact of Public Basic Research on Industrial Innovation: Evidence from the Pharmaceutical industry. Stanford, CA: Stanford Institute for Economic Policy Research.

Wagner, Caroline S. 1995. *Techniques and Methods for Assessing the International Standing of U.S. Science*. RAND MR-706.0-OSTP.

Watkins, Todd A. Forthcoming. Estimating Net Social Returns to a Case Study Cluster of ATP Photonics and Opto-Electronics Projects. Study underway in 2001 for NIST's ATP.

Werner, Oswald, and G. M. Schoepfle. 1987. *Systematic Fieldwork: Foundations of Ethnography and Interviewing*. Volume 1. Newbury Park: Sage Publications.

Yin, Robert K. 1984. *Case Study Research: Design and Methods*. 2nd Edition. Thousand Oaks, CA: Sage.

APPENDIX A. ELEVEN CASE STUDY METHODS: OVERVIEWS AND SUPPORTING MATERIALS

This Appendix presents brief overviews of the general characteristics and requirements of eleven methods that might be used in Office of Science case studies, ordered generally from the more qualitative to the more quantitative. Its purpose is to provide a brief introduction to a wide variety of methods, along with references and links to additional, more detailed information. This is not an exhaustive list: tools and methods that could be used in case study research are continually being developed, adapted, and refined. Most case studies use a combination of methods.

	Page
A1. Classic, Interview- and Records-Based Method	A-1
A2. Expert Review Method	A-5
A3. Process-Outcomes Method.....	A-9
A4. Historical Tracing Method	A-13
A5. Bibliometrics and Citation Methods	A-17
A6. Content Analysis Methods	A-23
(Including Co-Word, Data Tomography, Data Mining, and Visualization)	
A7. Sociometric/Social Network Methods	A-27
A8. User/Participant Survey Method	A-31
A9. Benefit-Cost (Cost-Benefit) Methods	A-35
A10. Statistical/Econometric Methods	A-41
A11. Key Indicator Method	A-45

The summaries provide the following information about each method, to provide program managers a grounding needed to plan and manage case study research:

- ◆ Overview
- ◆ Typical objectives
- ◆ Illustrative sources of data and methods of data collection
- ◆ Distinctive units of measure
- ◆ Resource issues
- ◆ Advantages
- ◆ Disadvantages
- ◆ Applications
- ◆ Suggested references.

This page intentionally left blank

A.1 CLASSIC, INTERVIEW- AND RECORDS-BASED METHOD

Overview

Narrative accounts of research and events surrounding and flowing from the research have value in generating human interest important to drawing a wider audience to the arena of scientific research. Narrative accounts can also promote understanding of complex processes by explaining the case in a story-telling format. Scientific research “makes a good story if it can be followed both as science and as narrative at the same time” (Gregory and Miller 1998:138). Using a “story-telling” approach, the writer/analyst provides a narrative that may include the genesis of the research ideas, an account of the main “actors,” the research goals and why they matter, the dynamics of the research, the results, and what came of them. Memorable stories of scientific research often will linger with an audience long after statistics have faded. They can add flesh to the bones of more academic and quantitative forms of evaluations and provide a broader, more holistic description of the case. Most case studies use this method to provide core information about the case being studied, supplementing it with other, more specialized methods to provide additional information about the particular topic of being studied.

Information for a narrative account is typically obtained from interviews with those involved or familiar with the case, and from documents, if available, that describe aspects of the case. Several interviewing techniques (snowball interviews, brainstorming, Nominal/Delphi, group interviews, impromptu interviews, etc.) can be used to capture the important elements of the case and the sequence of events that occurred (see Morgan (1997) for information on focus groups and Sudman and Bradburn (1982) and Werner and Schoepfle (1987) for guidance on interviewing). Source documents are particularly useful for establishing key dates, budgets, initial plans and goals, specific outputs, and other critical information helpful in framing the study.

Narrative approaches do not preclude the integrated use of models and statistics. In the process of telling the unique story of a research effort, the analyst may collect data on selected outputs, which, if compiled uniformly across a number of cases, can be used to provide not only project metrics, but also aggregate metrics for a portfolio of projects.

The test of a rigorous design for this method of study is whether it provides high internal and external validity. Roessner (2000) provides guidance on the design and analysis of studies that achieve this rigor. Eisenhardt (1989) discusses the use of the classic case study to build theories.

Typical Objectives

To interest, educate, and inform a larger audience, including non-scientists, about an area of research. To supplement quantitative program metrics. To describe the complex interactions among program elements and the broader institutional, social, and economic context.

Typical Sources of Data and Methods of Data Collection

Interviews usually conducted by analysts in-person or by phone
Surveys/questionnaires usually mailed or electronically administered
Project, program, or facility records, reports, and other documentation
News stories, press releases, and journal articles.

Distinctive Units of Measure

Text narrative, with or without supporting statistics, describing conditions, events, and developments, and summarizing patterns from interviews and secondary data sources.

Resource Issues

In addition to an analyst skilled in interviewing and records analysis, a skilled technical case writer who can make the story understandable and engaging is often important to the success of this method. Effective presentation means using the active voice and conversation-length sentences, including human-interest aspects of the story and addressing a real audience in appropriate language.

Advantages

This method is an effective way to reach a broad, diverse audience that may include policy makers, investors, potential users, educators, evaluators, and the general public. The time frame for the analysis tends to be quite flexible; the story may effectively focus on a single event or multiple events over a limited period of time, or it may cover a long period of time. This approach is appropriate for case studies of widely varying scope – from brief descriptive summaries to long complex analyses.

Disadvantages

Obtaining well-written, interesting cases for a broad audience is more difficult than is generally realized, and poorly researched and/or presented cases lose the audience and lack credibility. Although the scientific and academic community generally considers anecdotal evidence less persuasive than quantitative evidence, most people read and process anecdotal materials more easily. A good researcher, who is rigorous in data collection, analysis, and presentation, providing a variety of evidence and supplementing anecdotal descriptions with statistical results, can gain the advantages of both perspectives.

Examples of Applications of the Classic, Interview- and Records-Based Method

A common type of narrative, classical case study is one that looks retrospectively at the series of events that attributed to a particular, noteworthy outcome. These cases tend to focus on explaining how and why certain events contributed to the outcome, why they were important, and what lessons can be learned. For example, the National Academy of Science's Beyond Discovery™ website (www4.nas.edu/beyond/beyonddiscovery.nsf) provides some examples of how basic science research resulted in a variety of social benefits. Cases described range

from how global positioning systems (GPS) evolved from basic research on the atomic clock, to how an explosive, nitroglycerin was used to treat the heart condition angina. These examples provide public-friendly accounts of how basic research can help in unexpected and very rewarding ways.

Classic interview- and records-based case studies are also often used when analyzing unexpected or dramatic outcomes. These case studies can often take on an investigative tone. A case study was developed, for example, on the [arsenic contamination of the water supply for Bangladesh \(www.worldwaterday.org/cases/arsbgd.html\)](http://www.worldwaterday.org/cases/arsbgd.html) as a part of the United Nation's World Water Day 2001. This case study captures and presents in narrative form the essential factors attributing to the outcome, how the outcome was detected, and lessons learned from the event.

Many classical case studies will address less dramatic but equally important outcomes. The Federal Energy Management Program (FEMP) used classical, narrative case studies supplemented with some cost-benefit calculations (see section A.9) to describe how solar energy projects benefited federal facilities and reduced costs (www.eren.doe.gov/femp/millionroofs/ms-case.html).

James Gleick's *Chaos: Making a New Science* (1987) and Rhodes' (1986) *The Making of the Atomic Bomb* are examples of book-length case studies that use a combination of interviews, document review, and summaries of science to describe the development of key scientific concepts about complexity and dynamic systems and the development of nuclear weapon production capability. They illustrate the combined use of qualitative and quantitative data to tell engaging, yet rigorously researched stories.

An example of a collection of narrative case studies used to analyze a portfolio of projects is found in the Advanced Technology Program (ATP) of the Department of Commerce National Institute of Standards and Technology (NIST). Project narratives were developed for 50 completed ATP projects, supplemented by systematic data collection for a common set of variables related to (1) creation of scientific knowledge, (2) knowledge dissemination, and (3) commercialization of resulting technologies by U.S. companies. Rosalie Ruegg aggregated the data for each variable across the 50 projects to compute aggregate portfolio statistics. In addition, she used the data to develop composite performance scores for each project and showed the distribution of scores for the portfolio overall (Ruegg, 2001 National Academies of Science Report).

The prototype cases of Superconducting Wire and Magnet Technology at Brookhaven National Laboratory (Bozeman and Donez 1996) and Thermoacoustic Engines at Los Alamos National Laboratory (Klein and Bozeman 1998) developed for the R&D Value Mapping projects are also examples of this classic case study method. These studies were used to test the research methodologies and instruments to be used in a series of 30 case studies examining the economic and social impacts of research programs.

Suggested References

- Bozeman, B., and F. Donez. 1996. Brookhaven National Laboratory, Superconducting Materials and Magnet Technology. Prepared for Sandia National Laboratory. Atlanta, GA: Georgia Institute of Technology.
- Eisenhardt, K.M. 1989. Building Theories from Case Study Research. *Academy of Management Review* 14(4):532-550.

- Fiedler, J. 1978. *Field Research: A Manual for Logistics and Management of Scientific Studies in Natural Settings*. San Francisco: Jossey-Bass.
- Klein, H., and B. Bozeman. 1998. Los Alamos National Laboratory, Thermoacoustic Heat Engines. Atlanta, GA: Georgia Institute of Technology.
- Gleick, James. 1987. *Chaos: Making a New Science*. New York: Viking Press.
- Morgan, David L. 1997. *Focus Groups As Qualitative Research*, Second edition. (Qualitative Research Methods (Paper), Vol. 16). Newbury Park, CA: Sage Publications.
- National Academies of Sciences. *Beyond DiscoveryTM: The Path from Research to Human Benefit*. (Available URL: <http://www4.nas.edu/beyond/beyonddiscovery.nsf>. Last updated on 1/31/2001.).
- National Center for Case Study Teaching in Science: Case Study Collection. (Available URL: <http://ublib.buffalo.edu/libraries/projects/cases/ubcase.htm>. Last updated on 03/08/01.).
- Paradis, James G., and Muriel L. Zimmerman. 1997. *The MIT Guide to Science and Engineering Communication*. MIT Press.
- Roessner, David. 2000. Quantitative and Qualitative Methods and Measures in the Evaluation of Research. *Research Evaluation* 8(2):125-132.
- Rhodes, Richard. 1986. *The Making of the Atomic Bomb*. New York: Simon and Schuster.
- Ruegg, Rosalie. 2001. Taking a Step Back: An Early Results Overview of Fifty ATP Awards. *The Advanced Technology Program: Assessing Outcomes*. National Research Council/National Academies of Science. Washington, DC: National Academy Press.
- Sudman, S., and N.M. Bradburn. 1982. *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco, CA: Jossey-Bass.
- Yin, Robert K. 1984. *Case Study Research: Design and Tools*. Applied Social Research Methods Series Volume 5. Beverly Hills: Sage Publications.

A.2 EXPERT REVIEW METHOD

Overview

Expert review is judgmental assessment. Experts render their informed opinions orally or in writing, qualitatively or by numerical scores. The judgments are formed on the basis of the experts reviewing written or orally presented evidence or by direct observation. Experts may be assembled in conferring panels, functioning much as a jury, or they may perform their reviews independently, and at dispersed locations.

Expert reviews tend to be specifically focused, though they may address multiple cases. They are often used to provide an outside assessment of an existing research program or facility, as well as to provide advice on selecting projects, new directions for research, and technologies to fund. Most federal government agencies typically use several types of expert review methods. The Committee on Science, Engineering, and Public Policy (COSEPUP) describes three types of expert reviews: peer reviews, relevance review, and benchmarking. The following descriptions are taken from COSEPUP's *Evaluating Federal Research Programs* (1999):

Peer review – commonly used to make many kinds of judgments: about the careers of individual researchers, about the value of their publications, about the standing of research institutions, and about the allocation of funds to individuals and to fields of research.

Relevance review – used to judge whether an agency's research programs are relevant to its mission. Relevance review should not be confined to applied research, in which desired outcomes are defined. Relevance review should also consider basic research projects funded by federal agencies. Although the ultimate practical outcomes of basic research cannot be predicted, it is important to ascertain whether a given line of research is likely to contribute to an agency's mission.

Benchmarking – used to evaluate the relative international standing of U.S. research efforts. Although the principal reliance is on the judgment of experts, quantitative measures can also be used for confirmation.

Achieving a quality evaluation requires that the reviewers be highly knowledgeable about the subject; able clearly to articulate their opinions; free of conflict of interest; subject to a clear, rational, timely, and consistent process which leads all reviewers to apply evaluation criteria consistent with objectives; and provided anonymity if needed for objective, unbiased responses.

Typical Objectives

To assess quality of research; to make recommendations about allocation and reallocation of resources among programs, projects, and researchers; to provide program assessments for higher level administrative and policy decisions; and to generate project metrics, such as percentage of projects rated high, medium, and low in quality by a peer review process.

Typical Sources of Data and Methods of Data Collection

- Project or program goal statements, mission statements, progress reports, and other reports
- List of aspects to be assessed
- Evaluation criteria to be used
- Written and oral presentations by agency staff
- Direct inspection by reviewers
- Evaluation results and rankings from other studies.

Distinctive Units of Measure

Descriptive narratives and quality ratings, e.g., excellent/good/fair; high/ medium/low; satisfactory/unsatisfactory; numerical scores, e.g., 0-5.

Resource Issues

The main resource costs to implement this method are typically direct labor costs, plus travel costs, administrative costs, space costs, and possibly computer support costs.

Advantages

The method can provide a relatively quick, straightforward, and widely accepted approach to assessment. It is particularly useful for evaluating particular aspects of basic research programs where more objective methods may be impractical. While the method is suitable for assessing the quality of on-going research programs, it is less reliable for estimating their future impacts. It appears to be more reliable for identifying very good or very poor research than for dealing with in-between cases. It tends to bring in diverse points of view, and an interchange of ideas and feedback, which may alert managers to issues otherwise overlooked. It is widely respected and accepted as a fair, rational, and valid approach to program and project assessment.

Disadvantages

The quality of the review is dependent on the ability to identify qualified reviewers free of bias and conflict of interest. Perhaps the main disadvantage is that it is difficult to know the quality of the review or the soundness of decisions based on peer review since there is little empirical evidence regarding its reliability. Reports from expert reviews may provide limited descriptive information about the case and/or the evidence upon which the judgments were based.

Examples of Applications of Expert Review

The Department of Energy's Office of Science and Technology (OST) published a book (NRC 1999) on peer review: *Peer Review in Environmental Technology Development Programs*. This book describes how OST uses external experts to review their funded technologies at various stages of development. This review process uses a decision model to evaluate the technology maturation stages, from basic research all the way to implementation of technology. The review team compared the maturation of the science and technology to

“gates” that match the research to elements such as priority DOE needs and readiness for users. The review panel reports results back to program managers and others for feedback to decisions.

Another example is provided by Department of Commerce/NIST’s ATP, which uses a two-prong peer review process in selecting the technology projects it funds. Experts in the relevant areas of technology – for the most part scientists and engineers in government laboratories – provide written reviews and numerical ratings of proposals for research submitted by companies. Experts in business and economics primarily from the private sector, screened to avoid conflict of interest, provide written reviews and numerical ratings of business/economic plans contained in the research proposals. Both groups of experts are provided a list of ATP selection criteria and training prior to conducting their reviews and making their recommendations. The result of the peer review provides advice to the government selection board.

Suggested References

- Bozeman, B. 1993. Peer Review and Evaluation of R&D Impacts. In *Evaluating R&D Impacts: Methods and Practice*. Edited by B. Bozeman and J. Melkers. Pp. 79-98. Boston: Kluwer Academic Publishers.
- Committee on Science, Engineering, and Public Policy (COSEPUP). 1999. *Evaluating Federal Research Programs: Research and the Government Performance and Results Act*. Produced with support from National Academies of Science, National Academy of Engineering, Institute of Medicine. National Academy Press. (Available on-line at: www.nap.edu/catalog/6416.html).
- Geisler, E. 2000. The Metric of Peer Review. *The Metrics of Science and Technology*. Westport, CT: Quorum Books.
- Grol, R., and M. Lawrence (eds). 1995. *Quality Improvement by Peer Review*. New York: Oxford University Press.
- Kostoff, R. 1997. The Principles and Practices of Peer Review. *Science and Engineering Ethics* (Special Issue on Peer Review) 3:19-34.
- National Research Council (NRC), Committee on the Department of Energy-Office of Science and Technology's (OST) Peer Review Program. 1999. *Peer Review in Environmental Technology Development Programs*. National Academy Press. (Available on-line at: www.nap.edu/catalog/6408.html).

This page intentionally left blank

A.3 PROCESS OUTCOMES METHOD

Overview

A relatively new approach to describing and evaluating science and technology programs, proposed by Geisler and Rubenstein (2000), is the Process Outcomes Method. It brings a high degree of structure and comprehensiveness by providing descriptions of the flow of activities and outputs at each of the following five stages of the selected case:

- Stage 1 covers inputs to R&D, i.e., funding and other resources.
- Stage 2 addresses immediate outputs from R&D, such as publications, patents, software, methods, and theories.
- Stage 3 covers intermediate outputs, such as new and improved products and processes, or new methods of organizing.
- Stage 4 sees the generation of pre-ultimate S&T outputs, such as higher productivity rates in specific firms or sectors, cures for specific diseases, or reduced safety hazards for specific products or processes.
- Stage 5 yields ultimate S&T outputs in term of aggregate GNP effects, energy independence, quality of life improvements, or other broad socio-economic goals.

Geisler's and Rubenstein's process outcomes model also includes descriptions of the processes by which the S&T outputs at each stage of innovation are transformed and diffused to accomplish the desired end. Furthermore, for each stage, a set of "core indicators" is compiled (such as number of publications or number of awards received), and, additionally, a set of organization-specific indicators may be compiled (such as age profile of the scientists and engineers and their relative work experience).

As a final step, instead of performing traditional benefit-cost analysis, an attempt is made to cluster and index inputs and outputs. "Macro" indicators are constructed for each cluster of outputs at each stage in the innovation process. In addition, an overall, comprehensive index for all outputs may be constructed from the collection of indicators. Geisler terms the derived index of immediate outputs as the "alpha factor," which describes the outputs from the science segment of the S&T process. He terms the index of intermediate outputs, the "beta factor," which describes how companies and other organizations absorb and use the S&T outputs and transform them further. He terms the index of pre-ultimate outputs, the "gamma factor," and the index of ultimate outputs, the "omega factor," which indicates the total impacts of S&T on the economy and society.

Most other evaluation methods cover one or more of the stages covered by this method. For example, benefit-cost analysis draws investment costs from Stage 1 and benefits from Stage 5, and uses them to compute economic outcome measures such as net present value benefits, benefit-to-cost ratios, or rate-of-return on investment. As another example, the anecdotal, story-telling method may describe one, part of one, or all of the five stages of innovation. Furthermore, statistical methods are often used to collect and analyze core indicator data, and econometric analysis may be used to model relationships between inputs and outputs identified in the process dynamics. But only the process outcomes method covers all five stages and the related processes in a systematic way.

Typical Objectives

To provide a systematic evaluation framework that covers each stage, identifies the processes, relates processes to developments, and provides comprehensive performance measures.

Typical Sources of Data and Methods of Data Collection

Input identification (interview, project reports on budget and other resources)

Tiered outputs (interview, survey, project reports, databases)

Process dynamics (observation and analysis).

Distinctive Units of Measure

Immediate, intermediate, pre-ultimate, and ultimate output measures; indexes for each category of outputs; and an overall index for all outputs.

Advantages

It provides a comprehensive evaluative framework in that it lays out all the stages in the innovation process, together with underlying processes, and computes metrics for each stage. It can be used to address a single stage or some or all of the stages.

Disadvantages

There is not much guidance in actually computing the values, particularly the ultimate outputs (synonymous with outcomes). While the framework is useful, it glosses over the difficulties of obtaining all the output data suggested. The indices and factors used for metrics are not readily understandable without a tutorial, and it is not an approach that is familiar to most analysts. If sufficient data were available to flesh out the entire model, it should be possible to instead use more customary economic measures such as net benefits or return on investment; i.e., why use indices that are difficult to explain to most audiences, if the data are available to compute more familiar measures. In addition, there are methodological issues and difficulties entailed in aggregating diverse output data to construct indices and factors. Therefore, this method, in its entirety, is recommended for use only by analysts experienced in its use. Its principal use is as an organizing framework.

Examples of Applications of the Process Outcomes Method

Examples are rare, as the formal method is not widely used in its complete form. However, practically every evaluation employs many of its features. Its framework maps closely to the GPRA framework of inputs, outputs, and outcomes, and its attention to process dynamics supports management needs. Geisler (1995) uses the method to derive overall performance indices for two federal laboratories.

Suggested References

Geisler, E. 2000. The Metric of Process Outcomes. *The Metrics of Science and Technology*. Westport, CT: Quorum Books.

Geisler, E. 1995. An Integrated Cost-Performance Model of Research and Development Evaluation. *Omega* 23(3):Appendix A.

This page intentionally left blank

A.4 HISTORICAL TRACING METHOD

Overview

This method traces the evolution from research to a future significant outcome or from the outcome back to the research. It identifies critical antecedent developments, such as a developing pool of knowledge, and relates the developments to the ultimate significant outcome. The method may be employed in either of the following two ways:

- Forward Tracing, in which the analyst starts with the research of interest and traces forward the evolution of related subsequent events to an end point. The investigation determines where the research leads.
- Backward Tracing, in which the analyst starts with an outcome of interest and traces backwards to identify the critical stages or developments that were instrumental to the outcome. The investigation determines whether there is a path back to the research of interest and documents it if it is found.

Whether through forward or backward tracing, the historical tracing (also known as historiographic) method may use an interview/investigative approach to follow the trail from one organization or researcher to the next. Key events, people, documents, organizations are noted and the linkages among them established. Combining historical tracing with other methods, particularly bibliometrics and citation methods, may enhance the historical tracing method by identifying and providing details about linkages.

Rather than trace the evolution of a whole research program from input to output to outcome, an analyst may trace the evolution of some individual output of particular interest, or some phase of a program, or some field or topic. For example, suppose one of several objectives of a research program is to train scientists. A case study using the historical tracing method might identify graduate students who received assistance through the program's support of university research. The study could then trace the further development of the students into scientists, and, subsequently, identify their scientific contributions. Their connections to other scientists might also be traced. (Such a study might also be given as an example of the "Social Network Method.") As another example, the evolution of a field of research might be traced in a "genesis study," that did not trace linkages to downstream outcomes. Gleick's (1987) book tracing the development of chaos theory reflects this application of historical tracing.

Typical Objectives

To document how past investment led to and paid off in subsequent knowledge or technology. To develop better understanding of the evolutionary processes of science and technology. To suggest how future investment might pay off. To generate metrics of past performance.

Typical Sources of Data and Methods of Data Collection

Interviews

Search of historical documents, new articles, and journals

Linked information with dates or time sequences identified

Decision trees (applicable in prospective case studies)

Event trees (applicable in retrospective case studies)

Bibliometric and citation studies

Content analysis studies.

Distinctive Units of Measure

Identification of key developments and events in an evolutionary account; linkages and statistical measures showing strength of linkages.

Advantages

Such studies can provide interesting, memorable, historical accounts of evolutionary developments from research to ultimate outcomes. Because the studies are generally empirically based and verifiable, they tend to have high credibility.

- Specific Advantages of the Forward Tracing Approach: There is a well defined starting point for the analysis, and following related developments downstream preserves the linkage to the original research project.
- Specific Advantages of the Backward Tracing Approach: The significance of the outcome is already established rather than evolving.

Disadvantages

A long time period generally is covered with a complex chain of events; and, typically, multiple funding sources, multiple organizations, and many researchers are involved. The complexity makes it difficult to establish attribution among agencies, researchers, and research.

- Specific Disadvantages of the Forward Tracing Approach: If the project is still evolving, it may not lead to an outcome of interest, and the study may lead to a dead end.
- Specific Disadvantages of the Backward Tracing Approach: The targeted outcome may not be backward linked as expected to the research of interest.

Examples of Applications of the Historical Tracing Method Combined with Citation Analysis

CHI Research provides an example of a study that uses citations to trace the evolutionary linkage between NIH funding of ophthalmologic and related biomedical research and the subsequent development of patented eye-care technology. The study uses data on patents granted and the scientific literature cited by those patents, in a backward tracing study. It screens patents granted over a 20-year period from 1975 through 1994, using patent-office classifications and key words to identify all eye-care-related patents. It examines the patents' references to literature, and reviews references to scientific papers to determine the authors'

institutions and acknowledged funding sources. It finds that the National Eye Institute is the leading single institution in providing support for this research: 31% of all eye-care patents with science references cite papers that contain at least one acknowledgment to National Eye Institute (NEI) support; and when NEI is combined with the rest of the National Institutes of Health (NIH), 41% of the patents with science references are linked back to NIH-funded research. The study concludes that vision research sponsored by NIH is of direct and increasing relevance to the growing number of U.S. patented eye-care technologies. (The study can be found at the CHI Research Website: www.chiresearch.com). (Given its use of bibliometrics and citation methods, the study might also be given as an example of those methods.)

Other examples include:

- First TRACES Study—Backward tracing of selected major technological innovations to key events in their NSF-funded R&D history. (Conducted in 1967)
- Project Hindsight—Backward tracing of technology for 20 weapon systems funded by DOD through key events in their development. (Conducted in 1969)
- Second TRACES Study—Backward tracing of 10 innovations through the significant events in their NSF-funded R&D history. (Conducted in 1973)
- Subsequent TRACES Studies, e.g., for National Cancer Institute—Similar to the others except that additional features were added, such as citation analysis. (Conducted in 1989)
- DARPA Accomplishments Study— Study of 29 projects identified conditions for their success. (Conducted in 1991)
- DOE Office of Health and Environmental Research (OHER) Accomplishments Book —Description of the 40-year history of OHER, with selected accomplishments from different time periods highlighted and tracked. (Released in 1983 and 1986)
- DOE High Energy Physics Program —Description of history and qualitative assessment of accomplishments. (Conducted in 1990)
- The cases that provide the bases for the R&D Value Mapping cross-case analysis utilize tracing techniques (conducted in 1995-1998).

Suggested References

- Bozeman, B, J. Rogers, D. Roessner, H. Klein, and J. Park. 1998. *The R&D Value Mapping Project: Final Report*. Prepared for the U.S. Department of Energy, Office of Basic Energy Sciences. Atlanta, GA: Georgia Institute of Technology. (Available on-line at <http://rvm.pp.gatech.edu>).
- Kroll, P., G. Ault, and F. Narin. 1998. Tracing the Influence of Basic Scientific Research on Biotechnology Patents: A Case Study. *Patent World* (March):38-46.
- Maxwell, Robert A., and Shohreh Eckhardt. 1990. *Drug Discovery: A Case Book and Analysis*. Clifton, NJ: Humana Press.
- McMillan, G., F. Narin, and D. Deeds. 2000. An Analysis of the Critical Role of Public Science in Innovation: The Case of Biotechnology. *Research Policy* 29(January):1.

- Narin, F., K. Hamilton, and D. Olivastro. 1995. Linkage between Agency-Supported Research and Patented Industrial Technology. *Research Evaluation* 5(3 December):183-187.
- Perko, J., and F. Narin. 1997. The Transfer of Public Science to Patented Technology: A Case Study in Agricultural Science. *The Journal of Technology Transfer* 22(3): 65-72.

A.5 BIBLIOMETRICS AND CITATION METHODS

Overview

When researchers publish their theories and findings in archival journals, and when their filings for patents are granted, a massive database for S&T is created. Bibliometrics and citation analysis draw information from this database. Bibliometrics and citation analysis are used to assess the quantity, quality, significance, dissemination, and intellectual linkages of research. The methods are also used to measure the progress, dynamics, and evolution of scientific disciplines. The bibliometrics method and publications-citation analysis are particularly useful for assessing basic research because a typical output of basic research is publications. Patent-citation analysis, in contrast, is generally more important as a tool for assessing technology development, but may be useful in a historical tracing study of how basic research leads to downstream technology development. Both the bibliometrics and citations methods are often used in combination with other methods, such as with historical tracing and classic interview and records-based methods, to provide a richer assessment than any one of them alone will provide.

Bibliometrics can indicate the quantity of research by the counts of research outputs, such as articles in archival journals, peer-reviewed books, chapters in peer-reviewed books, reports, patents, keynote addresses at scientific meetings, and refereed conference proceedings.

Dividing numbers of outputs by research dollars, or some other input of interest, is a way to normalize the counts. The resulting ratio can serve as an indicator of research productivity – so many units of output per unit of input. For example, some companies and other organizations compare their average cost to produce a patent with that of other similar organizations. Counts of outputs and average number of a given output per unit of input are often used to track and analyze trends in science.

Analysis of the frequency with which others cite scientific outputs provides an indicator of the quality, significance, and dissemination of research. The more other scientists cite a research paper or report, the greater its assumed relevance, impact, quality, and dissemination, other things equal.

Tracking the volume of research outputs in a given field of science and their frequency of citation can reveal trends in the field and the rate of dissemination of knowledge in that field. Examining who is doing the citing can reveal where a field is moving. Investigating paper-to-paper, patent-to-patent, and patent-to-paper citations can help identify intellectual linkages among researchers and organizations producing papers and patents, as well as the knowledge linkages among subject areas. Clusters of relationships may be found. Citations of research papers in patents provide an indication of how research findings are being converted into technology.

The visual display of results is important to understanding and interpreting bibliometric results. It helps analysts and audiences comprehend the extent and direction of dissemination of research output.

Outputs are highly variable in type and quality. Normalization approaches can be used to allow cross-organizational or cross-discipline comparisons. One of several normalization approaches is to hold the journal constant and compare the number of citations a given paper, or group of papers, receives against the average citation rate of all papers in the journal (the

mean expected citation rate). A value greater than 1 indicates the paper, or set of papers, is more heavily cited than the average.

Bibliometric studies can be carried out on a macro-scale to characterize science activity at the national level. The Science and Engineering Indicators, for example, contain many bibliometric indicators at the national level. Studies can also be conducted at the micro-scale, to look at outputs of individual programs or of individual researchers.

Typical Objectives

To measure an important type of output of research programs. To investigate and indicate the significance of an organization's publications and patents. To investigate trends, emerging areas, and developing relationships in S&T research.

Kostoff (1997) cites the following objectives as appropriate for a suite of proposed bibliometric studies in large federally funded laboratories:

- To examine distribution of disciplines in co-authored papers, to see whether the multidisciplinary strengths of the lab are being utilized fully
- To examine distribution of organizations in co-authored papers, to determine the extent of lab collaboration with universities/industry/other labs and countries
- To examine the nature (basic/applied) of citing journals and other media (patents), to ascertain whether lab's products are reaching the intended customer(s)
- To determine whether the lab has its share of high impact (heavily cited) papers and patents, viewed by some analysts as a requirement for technical leadership
- To determine which countries are citing the lab's papers and patents, to see whether there is foreign exploitation of technology and in which disciplines
- To identify papers and patents cited by the lab's papers and patents, to ascertain the degree of the lab's exploitation of foreign and other domestic technology.

Kostoff also points out "while it was also recommended that the lab compare its output (papers/citations normalized over disciplines) with that of other similar institutions, this quantitative comparison should be approached with great caution." A comparative bibliometric analysis of 53 laboratories (Miller 1992 and 1996) clustered the labs into six types (Regulation and Control, Project Management, Science Frontier, Service, Devices, Survey). According to the study, "comparisons of scientific impacts should be made only with laboratories that are comparable in their primary task and research outputs."

Typical Sources of Data and Methods of Data Collection

Publications (publication of research results in refereed journals)

Peer Reviewed Books (research results published as commercial books reviewed by peers)

Keynote Addresses (invitations to deliver keynote addresses, or present refereed papers and other refereed presentations at major conferences related to one's profession)

Conference Proceedings (publication of research results in refereed conference proceedings)

Chapters in Books (research results published as chapters in commercial books reviewed by peers)

Competitive Grants (ability to attract competitive, peer reviewed grants from corporations and government agencies)

Patents (from the US Patent Office)
Existing publication and citation databases
Computer search engines.

Distinctive Units of Measure

Numbers of publications and patents; ratios of units of output to resource input; publication-to-publication citations, patent-to-patent citations, and patent-to-publication citations.

Resource Issues

Counting research outputs is generally best carried out by the research organization itself, by establishing internal output tracking systems. However, to perform bibliometric and citation analyses it may be helpful to utilize the services of organizations that have already developed large citation databases and search engines. CHI Research Inc., of Haddon Heights, NJ (www.chiresearch.com), performs publication-to-publication citation searches, patent-to-publication searches, and patent-to-patent searches for government clients that include NSF, NIH, and others. The Institute for Scientific Information (ISI) in Philadelphia (www.isinet.com) provides access to ISI citation databases covering over 8,000 international journals. It offers desktop access to cited references, provides users with updates of citation information, and offers training courses in the use of its databases. The U.S. Patent and Trademark Office (www.uspto.gov) offers on-line search capability for patent citations.

Advantages

Because knowledge is a principal output of research programs, the bibliometric and citation methods, with their focus on the documentation of knowledge, are widely used in evaluating S&T programs. They provide a means of assessing the quantity, significance, trends, and linkage of research based on fairly easy-to-collect output data. They offer more objectivity than many of the other methods. The methods scale easily without an explosion in costs. They can be applied at different levels, such as to analyze the multi-disciplinary strengths of a lab, the strength of its collaborations, and the strength of its research. They allow one to gain insight about international research directions by seeing which countries are citing certain papers and patents. They offer a relatively straightforward approach and produce measures easily understood by diverse audiences. Furthermore, these methods do not impose a burden on those evaluated. Also, since publications and citations tend to occur fairly close to the associated research, the time lag is much less for performing bibliometrics than the time required to measure more distant program outcomes. And it is generally accepted as a valid form of assessment.

Disadvantages

Publications and patents are only two types of research output and means of knowledge dissemination; others are ignored in bibliographic and citation analysis. Counts indicate quantity of output, not quality. One publication is not necessarily equal to another. Normalization approaches may not adequately adjust for differences in the quality and importance of journals. Differences in the publication or patenting rates of different organizations may reflect intended organizational publication practices rather than their different abilities to publish. Citations may not reflect a true intellectual linkage between the

citing source and referencing article. Works of poor quality may be heavily cited, particularly if they are controversial. Self-citations and friend-citations may bias and inflate citations rates. Inconsistencies in spellings, author names, and paper titles can cause citations to be missed in automated searches. There may be biases in the selection of key journals and in the process of selecting papers for publication. There may be problems in cross-disciplinary comparisons due to different publication practices, e.g., differences in the propensity of mathematicians to publish vs. the propensity of biotechnologists to publish, or due to differing stages of development in scientific fields, e.g., emerging fields versus mature fields. Furthermore, the resulting output measures do not prove that ultimate program outcome goals or benefits are achieved.

Examples of Application of the Bibliometrics and Patent Citation Methods

CHI Research studied the linkage between DOE/BES-supported research, as indicated by published papers and patents that cite the papers, to identify the contribution that the organization's research has made to U.S. patented technology. This and other examples of bibliometrics and patent citation studies can be found at CHI Research's website (www.chiresearch.com). As was indicated in section A.4, bibliometrics and citation analysis have been used in conjunction with the historical tracing method.

Suggested References

- Breitzman, A., and F. Narin. 1996. *Linkage between Patents and DOE BES Supported Papers*. CHI Contract No. 9521.
- Cozzens, S. 1989. *Literature-Based Data in Research Evaluation: A Manager's Guide to Bibliometrics*. Report to the National Science Foundation.
- Geisler, E. 2000. *The Metrics of Science and Technology*. Westport, CT: Quorum Books. (Can be ordered from the Greenwood Publishing Group at www.greenwood.com.)
- Kostoff, Ronald N. 1997. *The Handbook of Research Impact Assessment: Seventh Edition*. Defense Technical Information Center. Report No. ADA296021. (Available on-line at <http://www.dtic.mil/dtic/kostoff/handweb7index.htm>).
- Miller, R. 1992. The Influence of Primary Task on Research-and-Development Laboratory Evaluation - A Comparative Bibliometric Analysis. *R&D Management* 22(1):3-20.
- Miller, R. and A. Manseau. 1996. Bibliometric Indicators and the Competitive Environment of Research-and-Development Laboratories. *Scientometrics* 36(3):421-433.
- Narin, F. 1994. Patent Bibliometrics. *Scientometrics* 30(1):147-155. (Available on-line at CHI's website, given above)
- Narin, F. 1997. The Linkage between Patents and the Scientific Literature. Proceedings of a Conference on the Assessment and Use of Scientific and Technological Knowledge, entitled *Knowledge Production, Patents and Technological Intelligence*, Canberra, Australia. (Available on-line at CHI's website, given above)
- Popper, S. 1995. *Economic Approaches to Measuring the Performance and Benefits of Fundamental Science*. RAND PM-409-OSTP.

Wagner, Caroline S. 1995. *Techniques and Methods for Assessing the International Standing of U.S. Science*. RAND MR-706.0-OSTP.

This page intentionally left blank

A.6 CONTENT ANALYSIS METHODS

Overview

Several methods have been developed for extracting content information from science and technology texts and from interviews and other text-based sources. These methods are complementary to bibliometrics, adding content information to the quantity information provided by bibliometrics. They go beyond bibliometrics, in that they “extract patterns and relationships that measure the value of technological innovations” (Geisler 2000:188).

Co-word analysis uses key words (sometimes called “index words”) in searching text. The frequency of co-occurrence of the key words for a selected database of published articles depicts the evolution of ideas and concepts.

Database Tomography is a newer approach to co-word analysis that avoids the need to pre-specify key words. It can be applied to any text, including archival journal articles, as well as reports, memos, e-mail messages, etc. The texts to be searched are entered into computer storage and a computer-based algorithm extracts words and phrases that are repeated throughout the database, using the proximity of words and their frequency of co-occurrence to estimate the strength of their relationship. From the results, a map of the relationship among technical themes is constructed.

- Kostoff (1997) describes the process (paraphrased here):
Compute the frequencies of appearance in the text of all single words (for example, MATRIX), adjacent double words (METAL MATRIX), and adjacent triple words (METAL MATRIX COMPOSITES). Select the technical content words with the highest frequency as indicative of the pervasive themes of the full database (for example, SHOCK WAVE, REMOTE SENSING, IMAGE PROCESSING). For each theme word, compute the frequencies of words within ± 50 words of the theme word for every occurrence in the full text. Construct a word frequency dictionary that shows the words closely related to the theme word. Use numerical indices to quantify the strength of the relationships. Continue the process to arrive at final results, which identify (1) the pervasive themes of the database, (2) the relationship among these themes, and (3) the relationship of supporting “sub-thrust areas” to the high-frequency themes.

Textual Data Mining is a more recent development in content and pattern analysis that goes beyond statistical methods to employ such techniques as artificial neural networks and fuzzy logic to extract content information. Data mining computer methods employ rules and models of relationships that entail learning and use of the learning to predict how similar data sets will behave. Data mining is applied to “data warehouses” to discover the resident knowledge. Hence, there are close relationships among data mining, knowledge discovery, and knowledge management.

Techniques for Visualizing data bring to life the patterns and relationships identified by content analysis methods. Special software and hardware tools are needed to accomplish the visual displays. Examples of available visualization tools are SPIRE™ and Starlight.

Typical Objectives

To add content analysis to a bibliometric assessment. To map developments and emerging areas in science, and connections among areas. For military or strategic reasons to assess a nation's or organization's science focus and output. To extract patterns and relationships that indicate the value of knowledge or technology. To illustrate the historical evolution of research funded or conducted by a particular organization or set of organizations.

Typical Sources of Data and Methods of Data Collection

Existing databases

A selected group of documents entered into a computer database

Web-based, other electronically administered searches

Distinctive Units of Measures

Themes and sub-themes treated in textual material.

Advantages

Content-analysis methods provide more in-depth analysis of patterns, relationships among concepts, and relationships among disciplines than bibliometrics alone.

Disadvantages

Disadvantages specific to co-word analysis center on problems in selecting and using keywords, and on the inability to use full text of scientific documents. (The more recent content methods avoid these problems.) Different terms may be used to describe the same phenomena. Data mining and techniques for visualizing data are emerging fields in science and technology evaluation, and most evaluators are not experienced practitioners of these techniques. If data are not inclusive and/or not representative, the value of all these techniques is accordingly limited.

Examples of Application of Content-Analysis Methods

Kostoff (1997) gives the following example of database tomography applied to assessing a nation's science – in this case, the former Soviet Union (FSU):

Assume a paper represents about \$100K worth of effort. A 10,000-paper database represents about \$1B worth of effort, and comprises a representative sample of science output. A data tomography study of it would reveal much about the science interests, capability, and output of the originator.

Kostoff (1997) identifies the following additional applications:

- Identification of pervasive research thrusts in a database describing promising research opportunities for the Navy. The database consisted of thirty reports produced by the National Academy of Sciences panels and Office of Naval Research (ONR) internal experts on 15 technical disciplines
- Identification of pervasive thrusts in the 7400 project Industrial R&D (IR&D) database

- Identification of pervasive themes in a database whose narrative components describe each research project sponsored by the Department of Energy
- Identification of pervasive themes and their relationships in a database of journal articles consisting of one year's issues of the Journal of the American Chemical Society
- Identification of pervasive themes and their relationships in a database of journal articles related to Near-Earth Space Science and Technology.

Two examples of data visualization tools are SPIRE™ and Starlight, both developed by the Pacific Northwest National Laboratory. The Special Paradigm for Information Retrieval and Exploration (SPIRE) programs use computers and specialized methods of data visualization to analyze the content. The SPIRE programs read and summarize a document or set of documents, organize the information and create a map of the data on the screen. Information from thousands of different documents can be combined to generate patterns and relationships. The emerging result can often show unexpected clusters or other visual indicators of a pattern. Starlight is an advanced three-dimensional visualization technology used to help solve the problem of information overload. It has visual rendering capabilities similar to those described in SPIRE.

Suggested References

- Balch, P., and S. Brunak. 1998. *Bioinformatics: The Machine Learning Approach*, Cambridge, MA: MIT Press.
- Callon, M., J. Courtial, and F. Laville. 1991. Co-Word Analysis as a Tool for Describing the Network of Interactions Between Basic and Technological Research: The Case of Polymer Chemistry. *Scientometrics* 22(1):155-205.
- Geisler, E. 2000. *The Metrics of Science and Technology*. Westport, CT: Quorum Books. (Can be ordered from the Greenwood Publishing Group (www.greenwood.com)).
- Kostoff, R. 1997. *The Handbook of Research Impact Assessment*. Seventh edition. Office of Naval Research. Defense Technical Information Center, DTIC Order Number: A296021. (Available on-line at www.dtic.mil/dtic/kostoff).
- Risch, J.S., D.B. Rex, S.T. Dowson, T.B. Walters, R.A. May, and B.D. Moon. 1999. The STARLIGHT Information Visualization System. *Readings in Information Visualization*. Edited by S. Card, J. Mackinlay, and B. Shneiderman. Pp. 551-560. San Francisco: Morgan Kaufmann.
- SPIRE™ website: <http://www.pnl.gov/infoviz>.

This page intentionally left blank

A.7 SOCIOMETRIC / SOCIAL NETWORK METHODS

Overview

Scott (1991) identifies three types of social science data:

- Attribute data, for example, behaviors and attitudes of agents (analyzed using variable analysis)
- Relational data, for example, the contacts, ties, and connections that are properties not of agents but of systems of agents (analyzed using network analysis)
- Ideational data, for example, meanings, motives, definitions, and typifications (analyzed using typological analysis).

Network analysis deals primarily with relational data. Assessing communication linkages and exchange relationships among researchers is useful for suggesting the importance and sphere of influence of their work. It is a method that can be used in assessing basic science programs, where quantitative measurement of outputs and outcomes is more difficult. Interviewing, sending questionnaires, direct observation, and accessing secondary databases are techniques for obtaining the information on relationships needed to conduct sociometric/social network analyses.

One straightforward approach is to have researchers in the organization or discipline of focus list several researchers outside their organization with whom they most often share research information, and also several researchers they think are performing the most important work in the field. These researchers in turn are queried, and so forth. A multi-level communications network can be developed from the data. It can include the researchers' affiliations and disciplines. The network can reveal paths of knowledge spillover. It can show areas of influence and suggest the importance of the work of different researchers. And it can reveal both formal and informal collaborative relationships and the influence of one field on another.

Another approach to social network analysis is co-nomination analysis. A sample of researchers in a field is asked to nominate other researchers whose work is similar to or most relevant to their own. Based on the responses, networks are constructed. It is assumed that links exist between co-nominated researchers and that the strength of each link is proportional to the frequency of co-nomination.

A third approach, illustrated in Newman's (2001) analysis of scientific collaboration networks, analyzes data on co-authorship from databases such as MEDLINE and the Los Alamos e-Print Archive to examine the structure of scientific networks. Scientists are considered connected if they have authored a paper together (whether or not that paper is published in the peer-reviewed literature). This approach allowed analysis of a large (more than a million people) network without the requirement to collect primary data from the network participants. Network characteristics (such as number of authors, mean papers per author and authors per paper, number of collaborators, degrees of separation between scientists, and clustering) of the overall network and of different disciplines can be described and compared.

Researchers in the Institute for Policy Research and Development at the Georgia Institute of Technology have developed an approach they call Research Value Mapping (RVM) to assess the impact of research on human capital development. They map the career paths of

researchers involved in science and technology projects and facilities against features of these projects and facilities. They show how projects often lead to the development of different human resources and capacities, reflecting the non-linearity of the processes of research and innovation. With funding from NSF and DOE, and collaboration with researchers in France, the Institute is carrying out a number of case studies employing the RVM approach.

Typical Objectives

To assess the sphere of influence of researchers and the importance of their research. To identify pathways of knowledge spillover. To reveal collaborative relationships. To map human capital development from research projects.

Typical Sources of Data and Methods of Data Collection

Organizational charts

Lists of groups and individuals in groups (such as association membership; attendance at conferences, etc.)

Personnel lists

Interviews and questionnaires about relationships and networks

Databases that contain names (such as databases of publications)

Curriculum Vitae of scientists and engineers

Patent and publication citations

News releases

Articles, field reports, and proposals.

Distinctive Units of Measure

Relational networks of researchers. Descriptions of career paths and emerging capabilities.

Advantages

Data requirements can be relatively modest. The results are indicative of the role of a given institution and its researchers in advancing a field of science or contributing to human capital. The methods can be combined with any of the other evaluation methods to enrich the human dimension of analysis.

Disadvantages

Sociometric/social network methods may be less familiar to most government agencies and their stakeholders than some of the other methods. The analyses generally do not provide quantitative measures of value. It can be difficult to develop a clear definition of the boundaries that establish the network being described.

Examples of Application of Sociometric/Social Network Methods

Blau (1978) conducted a co-nomination study of theoretical high-energy physicists.

Garton, et al. (1997) studied social networks occurring on-line. They describe how to collect and analyze social network data, and demonstrate where social network data can be, and have been, used to study computer-mediated communication.

Newman (2001) studied the structure of scientific collaborations, as illustrated by co-authorship on scientific papers.

Rogers (undated) applied the RVM approach to analyze the human capital dimension of the University of California-Lawrence Berkeley Laboratory's Nuclear Magnetic Resonance Spectroscopy (NMR) program. The analyst traces over an extended period of time the career development and influence on other scientists of Dr. Alex Pines, the laboratory leader.

Suggested References

Blau, J. 1978. Sociometric Structure of a Scientific Discipline. In *Research in the Sociology of Knowledge, Sciences, and Art: An Annual Compilation of Research*. Edited by R. Jones. Greenwich: CT, JAI Press.

Freeman, Linton C. (undated). *Visualizing Social Networks*. (Available on-line at <http://tarski.ss.uci.edu/vis.html>).

Garton, Laura, Caroline Haythornthwaite, and Barry Wellman. 1997. Studying Online Social Networks. *Journal of Computer-Mediated Communication* 3 (1). (Available on-line at <http://www.ascusc.org/jcmc/vol3/issue1/garton.html>).

Georghiou, G., W. Giusti, H. Cameron, and M. Gibbons, M. 1988. The Use of Co-nomination Analysis in the Evaluation of Collaborative Research. In *Handbook of Quantitative Studies of Science and Technology*. Edited by A.F. J. Van Raan. North Holland.

Newman, M.E.J. 2001. The Structure of Scientific Collaboration Networks. *Proceedings of the National Academy of Science* 98(2):404-409.

Rogers, Juan D. Undated. University of California-Lawrence Berkeley Laboratory Nuclear Magnetic Resonance Spectroscopy; A Case Study of the R&D Value Mapping Project, Case 97-07, sponsored by the Department of Energy, Office of Basic Energy Sciences. Georgia Institute of Technology. (Available on-line at <http://rvm.pp.gatech.edu/cases/pines-case.html>).

Research Value Mapping (RVM) Program, Institute for Policy Research and Development School of Public Policy, Georgia Institute of Technology. Barry Bozeman, RVM Program Director. (See <http://rvm.pp.gatech.edu>).

Scott, John. 1991. *Social Network Analysis*. London: Sage Publications.

Social Networks Journal. Any issue. Amsterdam, The Netherlands: Elsevier Science Publishers.. (Available on-line at <http://eclectic.ss.uci.edu/~socnets/snjhome.html>).

This page intentionally left blank

A.8 SURVEY METHOD, INCLUDING USER AND PARTICIPANT SURVEYS

Overview

Asking individuals questions about their activities, knowledge, relationships, characteristics, and perceptions through the use of structured interviews or questionnaires is another method of evaluation and social research. Surveying users or customers of a program's research outputs, or users of a facility's service, or participants in a project or program to find out how they are affected and how they perceive performance and value is often a useful way to collect information for a case study. The survey can be administered in person, by telephone, by mail, or electronically.

Questions may be open-ended, requiring the analyst to decipher and classify responses in order to code them for statistical analysis, or they can be closed-ended for ease of tabulation. An example of an open-ended question: "What effect, if any, did using the research facility have on your organization's ability to meet its technology objectives?" An example of a closed-ended counterpart question: "How would you rate the effect of using the research facility on your organization's ability to meet its technology objectives? (a) highly significant, (b) moderately significant, (c) of little significance, (d) insignificant, or (e) don't know?"

A survey might, for example, focus directly on user satisfaction with a provided service or research output, on user opinion about the value of the service or research output, or it might question users about effects they experienced as a result of having the service or research result. Illustrative questions include: "How would you rate your satisfaction with the research facility? Were you: (a) completely satisfied, (b) somewhat satisfied, (d) somewhat dissatisfied, (e) completely dissatisfied?" "Would you rate the monetary value to your organization of using the research facility as approximately (a) =or>\$1 million, (b) =or>\$500 thousand but < \$1 million, (c) =or>\$100 thousand but < \$500 thousand, (d) =or >\$1,000 <\$100 thousand, (e) positive value, but unable to estimate in monetary terms, (f) no value, (g) don't know?" "As a result of using the research facility, would you say you were able to speed development of your technology by (a) 5 years or more, (b) 3 to <5 years, (c) 1 to < 3 years (d) < 1 year, (e) no noticeable change, (f) don't know?"

A user or customer survey can be administered on an ad hoc basis or it can be repeated routinely. It can be administered to control groups to provide a basis for comparison. Survey results can be compiled and tabulated, and aggregate and percentage responses reported. Statistical tests can be applied to the results to test significance.

The two key issues in conducting surveys are (1) sampling, i.e., describing the population and selecting the individuals to question; and (2) developing questions and response choices that are clear, unambiguous, and appropriate. The quality of survey results is highly dependent upon obtaining a sufficiently large and representative sample and asking them unbiased, revealing questions. Kish (1965) and Lohr (1999) provide clear discussions of sampling procedures. Dillman (2000) and Payne (1951) discuss the development of questions and their presentation.

Typical Objectives

To collect and analyze information from members of one or more groups, which will provide data on program performance. The results often can be used as performance indicators or measures. Comparisons can be made between groups with the sample. Relationships between attributes can be examined.

Typical Sources of Data and Methods of Data Collection

Mailed or electronically administered questionnaires

In-person or phone interviews.

Distinctive Units of Measure

Summary statistics, correlations, analysis of variance.

Resource Issues

A statistician may be needed to advise on drawing a sample from a population for study. Expert knowledge of survey techniques is also important to avoid bias in designing and administering a survey instrument. There are right ways and wrong ways to ask questions. Professional survey designers and administrators can help, either by designing the survey instrument, or by providing quality control to the design. There are organizations that specialize in administering surveys – private companies such as Ropers, Gallup, and Westat, Inc.; and certain university centers, such as those at Johns Hopkins University, the University of Baltimore, and the University of Michigan. Typical costs of surveys include cost of identifying the population to be surveyed, cost of designing and testing the survey instrument, cost of administering the survey, and cost of compiling and analyzing the data and reporting the results.

Government agencies are subject to the Paperwork Reduction Act and must obtain approval from the Office of Management and Budget (OMB) prior to administering most surveys. Many organizations also require review by a Human Subjects Committee before questionnaires can be administered.

Advantages

A survey can provide a relatively quick way to obtain measures of program effectiveness. Survey results are generally accessible to a broad audience. User surveys are a generally accepted method of evaluation. Surveys can provide information about participants and users not available through other sources.

Disadvantages

The importance of good survey instrument design, rigorous sample development, and diligent follow-up to increase response rates may be overlooked or ignored. Often those surveyed will need anonymity to prevent biasing the results, and this may restrict use and handling of the raw data, although it generally does not impede presentation of aggregate results.

Statistical results may be misused. Statistical reports of user surveys alone tend not to convey the richness and complexity of scientific research.

Examples of Application of the Survey Method

Sometimes certain aspects of a project are difficult to measure without surveys. INEEL and Idaho Asphalt Supply, Inc., for example, worked together to develop a biofiltration unit to reduce the odor associated with industrial exhaust. While certain gas emissions could be tested for reductions after use of the biofilter (sulfur emission were reduced by 99%), neighborhood surveys were better indicators of the odor issue. (See http://www.ornl.gov/news/pulse/pulse_v59_00.htm).

The Pacific Northwest National Laboratory, which operates the Environmental Molecular Science Laboratory for the DOE Office of Science, conducts an annual survey of facility users to assess user needs and satisfaction. (Results of the latest survey are available on-line at www.emsl.pnl.gov:2080/using-emsl/homepage.html).

Silber & Associates, an opinion research firm in Clarksville, Maryland, conducted a survey of ATP award recipients. Dr. Bohne Silber, an industrial-organizational psychologist and president of the firm, interviewed by telephone, using a structured interview guide, all 125 companies and consortia that participated in the ATP from 1990 to 1992. NIST contacted the ATP participating organizations in advance of the survey to notify them that Dr. Silber would contact them and to request their cooperation. Silber & Associates followed up with a phone call to schedule an interview with the lead researcher or manager in charge of the ATP effort. The survey instrument was pre-tested and revised several times before it was used. A copy of the report and the survey instrument can be found at ATP's website (www.atp.nist.gov/eao/eao_sil1.htm).

In another example from the ATP, Drs. Maryann Feldman and Maryellen Kelley (2001) conducted a survey that covered 502 proposals submitted in the 1998 ATP competition. The sample selected consisted of 100% of the winning firms and a random sample of 50% of the non-winners. The purpose of the survey was to collect information on the preparation of ATP proposals, the involvement of other organizations (formally or informally) in the ATP proposed project, the views of applicants concerning the fairness of the selection process, and experiences of winners and non-winners of awards since their ATP application. A point of interest was whether non-winners had been able to proceed with their proposed research plans, and whether winners and non-winner differed in their subsequent ability to attract other sources of funding. (The study reporting survey results is available on-line at ATP's website.).

Suggested References

Dillman, Don A. 2000. *Mail and Internet Surveys: The Tailored Design Method*. Second Edition. New York: John Wiley and Company.

Dillman, Don A. 1978. *Mail and Telephone Surveys*. New York: John Wiley and Sons.

Feldman, M., and M. Kelley. March 2001. Winning an Award from the Advanced Technology Program: Pursuing R&D Strategies in the Public Interest and Benefiting from a Halo Effect. National Institute of Standards and Technology, NISTIR 6577. (Available on-line at <http://www.atp.nist.gov/eao/ir-6577.pdf>).

Kish, Leslie. 1965. *Survey Sampling*. New York: John Wiley and Sons.

- Krishnaiah, P., and C. Rao. 1988. *Handbook of Statistics 6: Sampling*. North-Holland: Elsevier Science Publisher. (Abstract and ordering information for series available on-line at <http://www.elsevier.nl/inca/publications/store/5/2/4/3/8/9/>).
- Lohr, S.L. 1999. *Sampling: Design and Analysis*. Pacific Grove, CA: Duxbury Press.
- Payne, S.L. 1951. *The Art of Asking Questions*. Princeton, NJ: Princeton University Press.
- Sudman, S., and N.M. Bradburn. 1982. *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco, CA: Jossey-Bass.

A.9 BENEFIT-COST (COST-BENEFIT) METHODS

Overview

Benefit-Cost Analysis as used here encompasses a set of economic/financial approaches to quantify outcomes in relation to inputs. When used for private-firm financial analysis, the methods consider the net effect of all relevant positive and negative cash flows associated with a given investment decision that are captured or incurred by the investing firm. When used for federal public-sector analysis, the method takes a broader approach, including effects on society at large (usually defined as the nation) that can be monetized. Thus, the broad federal scope takes into account what would be “externalities” or “spillovers” to an investing firm – that is, effects resulting from an investment that are not incurred by or captured by the investor, but rather by others.

The economic/financial methods require quantifying effects in monetary terms. Consequently, the methods generally become more feasible the closer one moves from basic research to commercialization of new technologies. Because the set of related benefit-cost methods provide quantitative measures of outcome, rather than output, they are considered highly developed methods of evaluation.

Time Value of Money Adjustments: The methods take into account the monetized (dollar) value of benefits and costs that typically are spread unevenly over time. They adjust all amounts for differences in their times of occurrence to reflect the opportunity cost of capital and changes in the value of the dollar over time due either to inflation or deflation. This is called adjusting for the “time value of money.” Either the effects of inflation or deflation are first eliminated from the cash amounts so they are expressed in constant dollars (i.e., dollars with the same purchasing power) and a “real discount rate” is used to adjust for opportunity costs, or cash amounts are expressed in current dollars (i.e., dollars with uneven purchasing power) and a “nominal discount rate” is used to adjust for the combination of opportunity costs and inflation/deflation.

The discounting operation converts all dollar amounts to their equivalent values as of a common time so that they can be combined and compared with other discounted dollars. All amounts are expressed either as present values (a lump sum at the present), annual values spread evenly over a study period, or future values (a lump sum at a specified future date). Present values are most typically used in project/program evaluation.

The effect of discounting cash flows (i.e., streams of benefits and costs valued in dollars) over time is to reduce the value of amounts farther in the future relative to amounts occurring closer to the present time.⁸ An expanded set of discounting formulas, as well as multiplicative discount factors based on applying the formulas for \$1.00 of value, are readily available in most engineering economics or finance textbooks to cover the various discounting operations.

⁸ The basic equation for adjusting benefits occurring in a future year, t , to an equivalent amount occurring at the present is $B_t/(1+d)^t$, where B_t = benefits in future year t , and d = a discount rate. Thus, receiving benefits valued at \$100,000 in five years is equivalent to receiving \$71,299 today, if the discount rate is 7%. And, paying any more than \$71,299 today for a return of \$100,000 in five years would be a losing proposition since a 7% annual rate of return could otherwise be obtained.

A Set of Related Methods

Benefit-cost methods as defined here encompass several related methods that largely use the same input data, but provide different performance measures:

- *Net benefits*, also called *net present value*, is computed by subtracting time-adjusted costs (negative values) from time-adjusted benefits (positive values). A net benefit or net present value greater than zero indicates a worthwhile investment or project, since the minimum required rate of return is already accounted for through discounting. The measure is used for deciding whether or not to invest in a given project (invest if net benefits are positive); identifying the economic scale of investment (expand the investment as long as total net benefits continue to increase); selecting a portfolio of investments (select the combination for which aggregate net benefits are maximized); and assessing project or program impact in terms of the magnitude of dollars of net benefit produced by a project or program.
- *Life-Cycle Cost Analysis*, which can be used when the focus is on costs and changes in costs, sums all relevant costs associated with a specific investment choice or program alternative. A comparison of time-adjusted total costs among alternatives indicates which is lowest. Assuming equal levels of performance, the least-cost alternative is considered the cost-effective choice. The method can be used in evaluating programs or projects whose principal goal is to reduce the costs of a process, product, or service. For example, a project aimed at reducing operating costs of utility generation might be evaluated by comparing the projected total life-cycle costs of power generation without the project to actual or projected total life-cycle costs of power generation with the project. The difference in costs without the project minus costs with it provides a measure of project/program savings or loss.
- *Benefit-to-cost ratio*, a variation of which is *savings-to-investment ratio*, is computed by dividing benefits (or savings) by costs. The ratio indicates how many dollars of benefit per dollar of investment cost are realized. The ratio must be greater than one to indicate a minimally worthwhile project. Again, the minimal acceptable rate of return is already built into the analysis; such that a ratio of one means that the return is greater than the minimal acceptable rate. The ratio method can be used for evaluating project or program performance, but tends to be less used than net benefits for this purpose. A reason is that it does not show the magnitude of net benefits, a figure of keen interest in evaluation. Since a ratio computed on total benefits and costs begins to fall – as a project is expanded – before the optimal size is reached, it is important to compute ratios on marginal changes when using the method to size or scope a project or to allocate a budget among competing projects.
- *Internal rate of return* solves for the interest rate that will equate benefits and costs. (In contrast, the previous two measures specify the discount rate, which is expressed as an interest rate, and solve for the dollar amount or ratio, respectively.) For example, if one could spend \$71,299 today to receive \$100,000 in five years, it would be possible to use the discounting formula presented earlier to determine that the resulting rate of return is 7%. After the solution value of the interest rate is computed, it is compared against a specified minimum acceptable rate of return to determine the desirability of the investment or performance of a project. If one required, say, not 7%, but 10%, the above investment would not be attractive. An adjusted version of the measure, usually called “overall rate of return” or “adjusted internal rate of return,” avoids some inherent problems with the unadjusted version, such as the possibility of no unique solution value and the inherent assumption that the rate of return on the initial investment will also be obtained on reinvested

proceeds over the study period. Often studies that use economic/financial methods will include both rate-of-return (adjusted and unadjusted) and net-benefit measures.

- *Discounted Payback Period*, though not recommended as a sole performance measure, is included here because managers may encounter the measure and need to understand it. The method measures the elapsed time between the point of an initial investment and the point at which accumulated time-adjusted benefits or cost savings, net of other accumulated and time-adjusted costs, are sufficient to offset the initial investment. Thus, a project may be found to pay for itself in, say, five years, or eight years, or never. It is not recommended as a stand-alone measure because it does not provide an overall measure of return.

Benefit and Cost Estimation

The greatest challenge in using the benefit-cost method for evaluating federal program performance usually is estimating difficult-to-quantify benefits and difficult-to-identify costs. Typically the ideal data in the desired format are not readily available. Ingenuity on the part of the analyst is needed to use information that can be obtained at reasonable costs to estimate the values of interest. The analyst may use a variety of techniques to assist in the estimation. In conducting prospective studies, for example, analysts may find tools of decision analysis useful in estimating likely outcomes and associated values, diffusion models useful for generating market penetration of new technologies, and specialized tools useful for estimating particular kinds of benefits. For example, an analyst estimating the societal benefits of research leading to improved health might find useful the Quality-Adjusted-Life-Years technique to estimate the value to people of alternative health states. An application that requires placing monetary values on attributes of human well-being for which there are no observable market prices may be controversial, but the practice has considerable precedence in evaluation.

In other words, the skilled analyst may employ a variety of estimation techniques in conducting benefit-cost studies. There usually is no single way of carrying out this part of a study. The approaches and techniques used to estimate and monetize various benefits and costs will vary depending on the circumstances; the skills, experience, and judgment of the analyst; budgetary constraints; and the study objectives. David et al. (1992) provide a clear discussion of some of the issues associated with benefit and cost estimations related to basic science.

Dealing with Uncertainty and Risk in Benefit-Cost Studies

Applying any of the benefit-cost methods for project/program evaluation inherently entails working with uncertain or variable information. If the evaluation methods are used without incorporating techniques to account for uncertainty and risk, the results are stated as deterministic, quantitative answers and may be misleading in their implied level of precision. If probabilities can be attached to different values, risk assessment can be added to the benefit-cost analysis to indicate the extent to which the actual outcome will likely differ from the “best-guess” estimate. Risk assessment techniques include: Expected Value Analysis, Mean-Variance Criterion and Coefficient of Variation, Risk-adjusted Discount Rate Technique, Certainty Equivalent Technique, Simulation Analysis, and Decision Analysis.

Some of these techniques only account for risk exposure. Some account for both the risk preference of the science manager and the degree of risk exposure.

If probabilities cannot be obtained or estimated, a non-probability-based technique for treating uncertainty can be used. The principal of these techniques is Sensitivity Analysis, which tests how outcomes change as the values of input data about which there is uncertainty are changed. It shows the estimated outcome of a project/program for alternative data estimates and assumptions. It allows one to express the results in terms of a range of possible values. It reveals to the audience that there is uncertainty and indicates how the outcome might be affected.

Typical Objectives

To develop a quantitative estimate of outcomes, for evaluating performance of projects and programs.

Typical Sources of Data and Methods of Data Collection

Interview

Survey

Records, documents, and databases

Empirical observation of market-based data

Estimates of “shadow prices”

Econometric and statistical estimation models and techniques.

Distinctive Units of Measure

Net present value dollars

Benefit-to-cost ratio

Percentage annual rate of return on investment

Number of years required to pay back the investment.

Resource Issues

Economists generally design and perform benefit-cost studies of public programs. Financial analysts and economists generally apply the methods to evaluate decisions in for-profit companies. The evaluation of public programs is generally more difficult and is best performed by an analyst who is knowledgeable of spillovers and externalities and of methods of estimating them.

Advantages

Provides a quantitative measure of the impact of research in terms of ultimate economic outcome. Measures are stated in conventional financial terms, which facilitates understanding and comparisons. Quantitative measures of outcome, though they may be partial, generally provide more convincing evidence of value to some groups of stakeholders than the more qualitative measures.

Disadvantages

Because it is difficult or impossible to enumerate all the benefits from scientific research and estimate their value in monetary terms, only some of the benefits are usually captured by economic/financial measures. Knowledge-for-knowledge-sake benefits, for example, are excluded from consideration. Excellence in application of benefit-cost methods is highly dependent on the skills of the analyst in estimating difficult-to-measure benefits and costs. Because they focus on outcomes, the methods are often not applicable to the analysis of basic research programs, unless a historical tracing or process outcomes study has been done to link the basic research to downstream economic activities.

Examples of Application of Benefit-Cost Methods

The U.S. Department of Energy's Turbine Development Program needed to determine if funding two key wind turbine projects was worth the resources. A simple assessment of the costs to fund the Near-Term Product Development (NTPD) project (\$6.2 million) and the Value-Engineered Turbine (VET) project (\$2.7 million) were added to other incidental costs (testing, cost sharing by industry subcontractors), and the totals were compared to sales of turbines based on model designs generated through the program. Sales in 1996 totaled approximately \$330 million. Compared to an investment of \$12 million in research, this appears to be a good investment. Even a quick, back-of-the-envelope kind of benefit-cost assessment such as this can indicate the value of a program in a case study.

Economists at Research Triangle Institute's Center for Economic Studies performed seven case studies of ATP-funded applied research expected to lead to new medical treatments, using the Net Benefits and Internal Rate of Return Methods. The analysis included use of Decision Analysis, the Bass Diffusion Model, and the Quality-Adjusted-Life-Years technique in the estimation of benefits, and it employed sensitivity analysis to reflect uncertainties. It also included a counterfactual analysis in which outcome measures were computed two ways -- with the public investment and a counterfactual case that represented what the expected outcome would have been had there been no funding by the ATP. The goal was to estimate what part of the return was attributable to the federal program. Social rates of return, public rates of return, and private rates of return were calculated. A consulting economist, Prof. Albert Link of the University of North Carolina-Greensboro, performed an economic evaluation of radiopharmaceutical research at NIST that approximated first-level economic benefits in terms of the additional costs that radiopharmaceutical manufacturers and patients would likely have incurred in the absence of NIST's research. The study used a survey method for collecting information needed to estimate additional costs that would otherwise have been incurred, and calculated Net Benefits, Adjusted and Unadjusted Internal Rate of Return, and Benefit-Cost Ratio measures to evaluate the research program.

Suggested References

- Boardman, A. (ed.). 1996. *Cost Benefit Analysis: Concepts and Practice*. First edition. Englewood Cliffs, NJ: Prentice-Hall Inc. (Provides a textbook introduction to benefit-cost analysis through problem-solving exercises.)
- David, Paul A., David Mowery, and W. Edward Steinmueller. 1992. Analysing the Economic Payoffs of Basic Research. *Economics of Innovation and New Technology* 2(1):73-90.

- Gramlich, E. 1997. *A Guide to Benefit-Cost Analysis*. Second edition. Waveland Press.
(Provides a textbook treatment of benefit-cost methods.)
- Link, A. 1997. Economic Evaluation of Radiopharmaceutical Research at NIST. 97-2
Planning Report. (Available on-line at www.nist.gov/director/prog-ofc/report97-2.pdf).
- Mansfield, E. 1991. Social Returns from R&D Findings, Methods, and Limitations.
Research Technology Management November/December:24-27.
- Martin, et al. 1998. *A Framework for Estimating the National Economic Benefits of ATP
Funding of Medical Technologies*. Gaithersburg, MD: National Institute of Standards
and Technology, NIST GCR 97-737. Prepared by Research Triangle Institute, Center for
Economic Research. (Available on-line at www.atp.nist.gov/eao/eao_pubs.htm).
- Ruegg, R. 1997. Economic Methods. In *CRC Handbook of Energy Efficiency*. Edited by F.
Kreith and R. West. Boca Raton: CRC Press. (Provides a survey of benefit-cost methods
and methods for treating risk and uncertainty.)
- Ruegg, R., and H. Marshall. 1990. *Building Economics: Theory and Practice*. New York:
Van Nostrand Reinhold. (Provides a detailed practical guide to structuring and carrying
out benefit-cost case studies, with many examples based on energy systems in buildings.

A.10 STATISTICAL/ECONOMETRIC METHODS

Overview

Statistical, mathematical economics, and econometric methods are useful for analyzing functional relationships between economic and social phenomena, such as inputs and outputs. Econometrics uses statistical and mathematical analysis to analyze economic relationships. In case studies, statistics are used to describe and compare subgroups within a case as well as to describe and compare cases (in multi-case studies). Attributes of the cases studied may also be compared to those developed from broader econometric studies or surveys.

Statistical Methods: A variety of statistical methods are used in evaluation, a few of which are briefly treated here:

- *Descriptive statistics* provides tabular, graphical, and numerical summaries of data. The purpose of descriptive statistics is to facilitate the presentation and interpretation of data, for example, for characterizing research programs, their outputs, and their effects.
- *Statistical inference*, the process of using sample data to make inferences about a population, is often used in evaluation studies using survey techniques to reduce the time and cost of collecting data from the entire population.
- *Hypothesis testing*, a form of statistical inference that uses data from a sample to draw conclusions about a population, first makes a tentative assumption called the null hypothesis, denoted by H_0 . Then an alternative hypothesis, denoted H_a , the opposite of what is stated in the null hypothesis, is defined. The hypothesis-testing procedure generally uses sample data to determine whether or not H_0 can be rejected. If H_0 is rejected, the statistical conclusion is that the alternative hypothesis H_a cannot be rejected. For example, H_0 may be that basic physics research is not directly linked to achieving biological breakthroughs, and H_a , the opposite.
- *Regression analysis* involves identifying the relationship between a dependent variable and one or more independent variables, such as the relationship between numbers of publications, the field of science, and the size of the research budget. A model of the relationship is hypothesized, and estimates of the parametric values are used to estimate a regression equation. The estimated regression equation is then used to predict the value of the dependent variable given values for the independent variables.
- *Correlation analysis* measures the degree of association between variables. But neither regression nor correlation analyses prove cause-and-effect relationships. They can indicate only how or to what extent variables are associated with each other.

Econometric Methods:

- *Econometric analysis* works with production functions, as well as cost functions and supply functions. The production function is a mathematical expression of the technical relationship between inputs (or factors of production) and the outputs obtained. The equation states the amount of output that can be obtained from combinations of factors, assuming that the most efficient available methods of production are used. The production function can be used to answer a variety of questions, such as the change in output from an additional unit of a particular factor

of production, or the least-cost combination of productive factors that can be used to produce a given output. Evaluation examples using econometric methods include assessing the impact of government R&D funding of industry research on company performance and productivity. Another example would be analyzing the mechanisms and instruments of public research programs to find how they relate to program objectives.

- *Macroeconomic modeling* entails using mathematics and statistics to describe an entire economy or major parts of it. Macroeconomic models are used for economic forecasting and the analysis of public policy, such as analysis of S&T policies aimed at economic growth. An example of a macroeconomic model that is sometimes used in evaluation is the REMI Policy Insight™ forecasting and policy analysis model. It has been used to forecast the economic effects of a wide range of policy initiatives for purposes of evaluating the policies, and to forecast the regional or national economic effects of public- and private-sector investments.

Typical Objectives

To estimate outputs from inputs. To test for relationships between inputs, program mechanisms, and outputs. To assess policy options.

Typical Sources of Data and Methods of Data Collection

Existing public- and private-sector databases, such as the Corptech database, patent databases, census data, tax records, input-output data tables, labor data, and prices.

Survey

Interview

Empirical observation.

Distinctive Units of Measure

Quantitative results and indicators, the units of which depend on the nature of the analysis.

Advantages

Statistical and econometric methods produce quantitative results with detailed parameters. They can contribute to understanding relationships between inputs and outputs. They increase the analytic capability of evaluators and can provide standard measures of significance, and uncertainty.

Disadvantages

Analyses and results may be difficult for the non-specialist to understand, replicate, and communicate. The quality of the results is dependent upon the quality of the sample and validity of the measures. Not all effects of scientific research can be captured in highly quantitative methods or in monetary terms. Technical difficulties arise in including technical information in production function analyses. The processes through which technical knowledge impacts the larger economy are not well understood, nor well captured in available models.

Examples of Applications of Statistical and Econometric Methods

The great diversity in application of these methods is illustrated by the following examples taken from the ATP, which has used econometric methods more than most other government programs.

Maryann Feldman and Mary Ellen Kelley (2001) used survey data, sampling techniques, and multiple statistical regression modeling methods to estimate the contribution of ATP awards to selected indicators of success. They employed controls for differences in organizational resources, capabilities, prior experience, and types of R&D projects.

Nicholas Vonortas (ATP publication expected in 2001) analyzed two databases – one from the U.S. Department of Justice’s National Cooperative Research Act and one from the Advanced Technology Program – to compare ATP joint-venture formations with non-ATP joint venture to test the hypothesis that the collaborations ATP thinks it is fostering would have occurred anyway. He concluded that the hypothesis did not hold, and ATP was fostering the formation of joint ventures.

Mark Ehlen (1999) used the REMI Model in an economic evaluation case study of research leading to a new automotive process technology. In addition to estimating effects of the government investment on Gross Domestic Product (GDP), Ehlen applied the model to estimate employment and income tax effects.

Lee Branstetter and Mariko Sakakibara (ATP publication expected in 2001) used econometric methods to assess the impact of ATP research joint ventures on R&D productivity of project participants. The study used patents granted as the measure of R&D output. The study, using limited data, suggests that ATP-funded research joint ventures have a positive impact on firm patenting.

David Austin and Molly Macauley (2000) developed a cost index model for estimating expected consumer welfare gains from substantial improvements in existing products resulting from ATP-funded research. They tested their model by estimating effects of two innovations in digital data storage.

Suggested References

Austin, David, and Molly Macaulay. 2000. Estimating Future Consumer Benefits from ATP-Funded Innovation: The Case of Digital Data Storage. NIST GCR 00-790. Washington, DC: Resources for the Future. (Available on-line at http://www.atp.nist.gov/eao/eao_pubs.htm).

Blalock, H. M., Jr. 1972. *Social Statistics*. Second edition. New York: McGraw-Hill.

Ehlen, Mark A. 1999. Economic Impacts of Flow-Control Machining Technologies: Early Applications in the Automotive Industry. NISTIR 6373. Gaithersburg, MD: NIST. (Available on-line at http://www.atp.nist.gov/eao/eao_pubs.htm).

Mansfield, E. 1991. Academic Research and Industrial Innovation. *Research Policy* 20(1):1-12.

Policy Analysis Applications of REMI Economic Forecasting and Simulation Models. 1995. *International Journal of Public Administration* 18(1):13-42. (Provides an overview of the REMI Economic Forecasting and Simulation Models used for policy analysis.)

Popper, S. 1995. Economic Approaches to Measuring the Performance and Benefits of Fundamental Science. RAND PM-409-OSTP.

The ATP studies described under “Applications” are either available now on-line at http://www.atp.nist.gov/eao/eao_pubs.htm, or are expected to become posted at this site.

A.11 KEY INDICATOR METHOD

Overview

Indicators point towards, suggest, or provide evidence of a likely outcome. Indicators are often used as interim evaluation results for applied research projects that are currently underway. They sometimes are used as the sole results when it is too costly or difficult to get at the actual measures of outcome.

The data used for indicators often come from application of the other methods of evaluation, such as survey, bibliometrics, and statistical methods. The unique aspect of the method is the process of constructing measures that will serve as good indicators. Indicators may be presented as annual data or cumulative trend data, adjusted or unadjusted for costs. Examples of indicators are number of annual awards received from outside organizations; cumulative number of graduate students supported; and average cost (constant dollars) per peer-reviewed publication.

Research Inputs (Costs; Expenditures) as a Performance Indicator – Macro S&T indices, such as those reported biennially in *Science & Engineering Indicators* report aggregate national research expenditures as indicators of the nation's S&T effort. The premise is that aggregate expenditures are useful indicators of the socio-economic outcomes of science, technology, or research strength. Costs are generally much easier to capture and are already expressed in dollars, making them an attractive proxy for the more difficult-to-measure outputs and outcomes and a convenient way to track and compare the magnitude of effort among programs, organizations, sectors, and countries.

Input data can be used to construct various index measures, such as a measure of a nation's R&D intensity, computed as R&D expenditures divided by GNP. R&D input data and indexes constructed from them are best used for constructing broad indicators of R&D magnitude or intensity, e.g., for indicating the strength of a country's science and technology system or indicating the direction of innovation.

While research cost is generally accepted as a legitimate and useful indicator of resource commitment, it is not considered an appropriate project or program performance indicator. Using inputs as a proxy for outputs, or applying multipliers to costs and portraying the results as a comprehensive benefits measure are practices that generally lack credibility in the evaluation community. Research cost is an input measure, not a measure of output, outcome, or performance. The size of the research expenditure does not show what the resources were used to produce, how effectively and efficiently they were used in production, what additional resources were required, or the value of what was produced.

Outputs as a Performance Indicator – Outputs are frequently reported in performance measurement. They offer the typical advantages of being directly observable in the shorter run, easier to measure than longer-term outcomes, generally of interest to stakeholders, and providing evidence that developments are underway which are expected to lead to ultimate goals. A single output or a set of outputs may serve as an indicator of performance.

Examples of Output Measures used as Indicators of Desired Outcome

- **Desired Outcome:** significant contributions to the scientific knowledge base. **Output Indicator:** awards received for scientific excellence or achievements.

- Desired Outcome: enhanced research capability in a specific field.
Output Indicator: number of post-docs supported in the field.
- Desired Outcome: broad dissemination of research results.
Output Indicators: papers published, patents granted, and citations of both.
- Desired Outcome: research results lead to improved health of citizens.
Output Indicator: number of derived medical procedures in clinical trials.

Typical Objectives

Input data are used to compare the magnitudes of and trends in R&D programs, organizations, sectors, and countries. *Output data* are often used as performance indicators when (1) a project is still on-going and outcomes have not yet occurred; (2) outcomes are too difficult or expensive to measure, but certain outputs can feasibly be measured; and (3) to supplement measures of outcomes and provide more detail about how they occur.

Typical Sources of Data and Methods of Data Collection

Budgets
Expense reports
Cost tracking software
Other studies, such as surveys
Program documents
Empirical observation.

Distinctive Units of Measure

Dollars of costs; counts of outputs; trends in outputs.

Examples of Application of Key Indicator Analysis

R&D expenditures in the U.S. versus that in other countries is reported by the U.S. National Science Board as an indicator of comparative S&T strengths. (See, for example, Fig 2-27 in *Science & Engineering Indicators 2000*).

Some research-intensive companies use R&D cost per patent granted as an indicator of research efficiency.

Output data is often used by agencies for GPRA reporting. See the annual performance metrics reports of agencies.

Suggested References

National Science Board. 2000. *Science & Engineering Indicators*, Washington, DC: U.S. Printing Office. June. (Available on-line at www.nsf.gov/sbe/srs/seind00/start.htm). (Provides U.S. R&D and S&T input data used as indicators.)

OECD. 2001. *Main Science and Technology Indicators 2000*. Provides the most frequently used indicators on scientific and technological performance of OECD Member countries, including 89 indicators. (Available on-line at www.sourceoecd.org/bookshop).

Agency Annual Performance Plans detailed in fiscal year Congressional budget requests includes information on agency performance goals and targets, verification and validation data, and actual measures of projected performance, which often include output measures as indicators. These reports are available from various agencies of government.

Powell, J. 1999. *Business Planning and Progress of Small Firms Engaged in Technology Development through the Advanced Technology Program*. National Institute of Standards and Technology, NIST 6375. Provides an example of output data collected by electronically administered survey to indicate how small businesses are performing in the Advanced Technology Program. (Available on-line at www.atp.nist.gov/eao/ir-6375.pdf).