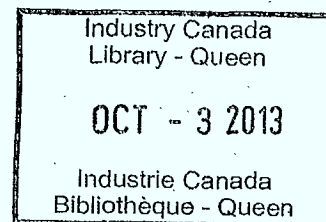# Final Report
Methods for Assessing the Socioeconomic Impacts of Government S&T

ARA

FINAL REPORT
METHODS FOR ASSESSING THE SOCIOECONOMIC IMPACTS OF GOVERNMENT S&T

Sponsoring Departments and Agencies:

Agriculture Canada
Atlantic Canada Opportunities Agency
Canada Centre for Mineral and Energy Technology
Environment Canada
Fisheries and Oceans
Geological Survey of Canada
Industry, Science and Technology
National Defence
National Research Council of Canada
Office of the Comptroller General
Supply and Services Canada
Transport Canada

The ARA Consulting Group Inc.

May, 1993

Table of Contents

Executive Summary

This study was commissioned by an interdepartmental committee of ADMs to identify the best methods available for the federal government to determine the economic and social impacts of its science and technology (S&T) activities. The main study activities were:

- a review of the literature dealing with appropriate methodologies for S&T impact assessment;

- a review of the literature describing current practices in other jurisdictions (e.g., other countries);

- an overview of the current practices of the Canadian federal government in this area.

Most of the data collection dealt with research and development (R&D) activities, and, therefore, in the remainder of this summary we refer to R&D rather than S&T.

The main conclusion that was drawn from the review of impact assessment methods is that there are methods available which can provide at least a partial picture of the socioeconomic impacts of government R&D activities and, in many cases, can provide reasonably accurate quantitative estimates. The main factors that determine which methods are most appropriate are:

- the timeframe for the assessment—i.e., whether the impacts of an R&D activity are being assessed prior to or after the R&D has been completed;

- the type of the R&D—basic/strategic research, applied research, or product/process development;

- the purpose of the R&D.

Exhibit E.1 summarizes the applicability and the most important characteristics of the impact assessment methods which are the most useful for assessing the impacts of government R&D.

Other important findings from this study are as follows:

- There is an increasing emphasis on assessing the impacts of government R&D, both in other countries and in Canada.

- There are two significant changes occurring within the impact assessment practices of a number of countries—1) there is a increasing emphasis on prospective impact assessment, and 2) there is an increasing emphasis on more careful consideration of the "receptor capacity" (level of interest and capability) of the organizations which are intended to be the primary users of the R&D.

Exhibit E.1: Applicability and Characteristics of Different Methods

Methods	R&D Time Frame	R&D Type	R&D Purpose	Strengths and Weaknesses	Relative Cost
Modified Peer Review	Past, on-going, and future	All	All - although least well suited for Category 1	**Strengths** • Relatively easy to organize. • Can provide valuable information on potential impacts • Probably the best method for basic/strategic R&D. **Weaknesses** • Relies on the opinions of a small number of people. • Qualitative information only.	Low/Medium
User Surveys	Past and on-going	Applied research and development	Category 2 to Category 4	**Strengths** • Overcomes the problem of a small number of respondents. • Possible to develop quantitative indices. **Weaknesses** • Structuring the survey and analyzing the results can be tricky.	Medium (often requires considerable time to identify users, develop survey methodology, and analyze results)
Benefit-Cost Methods	Past (can be used for on-going and future R&D in certain circumstances)	Applied research and development	Category 4 (can be used for Category 2 and Category 3 R&D in certain circumstances)	**Strengths** • Can provide reasonable and defensible estimates of potential benefits. • Provides a structured framework for assessing R&D projects which forces the right questions to be asked. **Weaknesses** • Can be very time consuming and labour intensive. • Results are critically dependent on assumptions which can be highly uncertain. • Because of cost and time requirements can only be used for a limited number of projects.	High (data collection requirements are very demanding)

R&D types: Basic/strategic research, applied research, product/process development
R&D purposes: Category 1 - R&D infrastructure, Category 2 - Policy development, Category 3 - Policy attainment, Category 4 - Industrial innovation

Exhibit E.1: Applicability and Characteristics of Different Methods

Methods	R&D Time Frame	R&D Type	R&D Purpose	Strengths and Weaknesses	Relative Cost
Case studies	Past	Applied research and development	Category 2 to Category 4	**Strengths** • Can provide good illustrations of the relationship between R&D and its impacts. • Probably the best method, in general, for Category 2 R&D. **Weaknesses** • Generally there is no way to "add up" the results of a group of case studies to obtain a measure of the total impacts of the group. • The results cannot be extrapolated to other R&D projects which are not in the group.	Medium (depending on the number of case studies)
Partial Indicators	Past and on-going (and future to a limited extent)	All	All	**Strengths** • The information required to specify the indicators is relatively easy to collect. • Probably the best method for on-going monitoring. **Weaknesses** • The individual indicators can generally only be "added up" on a judgemental basis, making overall impact assessment more difficult. • Provides only a very partial picture of impacts.	Low
Integrated Partial Indicators	Future	Applied research and development	Category 2 to Category 4	**Strengths** • An easy but structured of way to identify research priorities. • Forces the decision makers to explicitly consider the key determinants of impacts. **Weaknesses** • Totally relies on the judgement of (usually a few) individuals. • Potential for bias in assigning weights to different criteria.	Low

R&D types: Basic/strategic research, applied research, product/process development
R&D purposes: Category 1 - R&D infrastructure, Category 2 - Policy development, Category 3 - Policy attainment, Category 4 - Industrial innovation

- With regard to the sufficiency of the information which is currently available to Canadian government departments, departments interviewed generally feel that there is a need for more information on the impacts of R&D, especially regarding the potential impacts of on-going and future projects.

The study also included a preliminary assessment of several potential problems which might inhibit the practice of impact assessment within the Canadian government. The findings were as follows:

- In general, the level of interest within Canadian government departments and agencies in assessing the impacts or potential impacts of their R&D activities does not seem to pose a problem—i.e., the level of interest is high.

- The capability to apply the full range of impact assessment methodologies internally, on the other hand, may be a problem in some departments , but this is relatively readily solvable.

- Getting researchers and research managers to collect the data necessary for impact assessment and to document the impacts of R&D is a definite challenge.

- Another challenge is the lack of uniform knowledge regarding impact assessment methods.

In spite of the challenges noted above, it can be concluded from this study that it is both feasible and useful to incorporate socioeconomic impact assessment as an element of management practice for government R&D activities. It needs to be kept in mind, however, that the appropriate amount and intensity of impact analysis, as well as the methods used, will vary with the type and purpose of the R&D and the information needs of the department or agency.

1.0 Introduction

1.1 Study Purpose and Rationale

This study was commissioned by an interdepartmental committee of ADMs to identify the best methods available for the federal government to determine the economic and social impacts of its science and technology (S&T) activities. The rationale for this study, as quoted from the request for proposal (RFP), is as follows:

> *In recent years the R&D[1] environment both in Canada and abroad has changed drastically. There are an increasing number of research opportunities and needs coupled with restricted funds available for R&D activities. Internationally, competition among countries has forced governments and corporations to make strategic decisions about their R&D investments based on a clear understanding of the value, impacts, and benefits of those expenditures. In this context, decision makers need tools which support good planning and monitoring of their resources.*

The RFP goes on to note that strategic decision making requires this kind of information and, in addition, S&T program managers need information on the social and economic impacts of their activities to monitor the performance of their programs.

The need for federal S&T departments and agencies to know about and be able to apply the best available methods for impact assessment is very real and a matter of some urgency. This need has been primarily brought about, of course, by the emphasis on deficit reduction and the consequent pressure on all government activities to be able to demonstrate their value. Both central agencies and external "clients" are asking about the impacts of S&T activities on an increasingly frequent basis. For example, in recent years, there has been greatly increased pressure exerted by central agencies to carry out formal evaluations of S&T programs. Many of these programs had successfully resisted evaluation by using the argument that they were unevaluable. However, this argument is no longer valid because of the large number of successful evaluations carried out in the late 1980's and early 1990's, and because the methods used and results achieved in these evaluations have gained widespread acceptance.

S&T departments and agencies that have not been actively using the latest impact assessment methodologies are subject not only to demands by central agencies but also to the consequences of erroneous conclusions which may be drawn by their clients—conclusions which, due to the lack of documentation on impacts, may not always be based on accurate information. In one recent example an industry review committee criticized the relevance of the research being

---

[1] The original focus of this study was on R&D activities, not the somewhat broader scope of all S&T activities. As discussed in Section 1.3, this report will focus primarily on R&D activities.

carried out by a particular group of government labs to the needs of industry. The labs felt that the committee's conclusions were based on inaccurate information. However, in the absence of defensible documentation on the impacts of their research these conclusions were difficult to refute.

The need for information on the economic and social impacts of S&T activities is not just limited to activities that have been carried out in the past. Senior departmental officials are beginning to require that the managers of R&D programs include criteria related to potential economic and social impacts in assigning priorities, allocating resources, and deciding which R&D projects to include in their work plan.

This study is intended to provide information that will be useful to the S&T departments in planning and carrying out these impact assessments.

1.2   Work Plan

The terms of reference for this study called for the identification of the best methods for assessing the economic and social impacts of S&T activities to be based on:

- A review of the literature dealing with appropriate methodologies for S&T impact assessment.

- A review of the literature describing current practices in other jurisdictions (e.g., other countries).

- A review of the current practices of the Canadian federal government in this area.

The terms of reference also required that the study take into consideration the information that is needed by S&T program managers and senior officials in S&T-based departments for strategic decision making, planning, and review purposes.

The study can conceptually be illustrated as shown below.

```
(1) Theory re. methods for        (2) Current state of practice      (3) Current state of practice
    assessing the impacts and         in other organizations             in the Canadian
    benefits of S&T                                                      federal government


                    (5) Possible methods and          (4) Information needed by
                        their pros and cons                federal S&T
                                                           decision makers


                            Identification of best methods for
                                 federal S&T activities
```
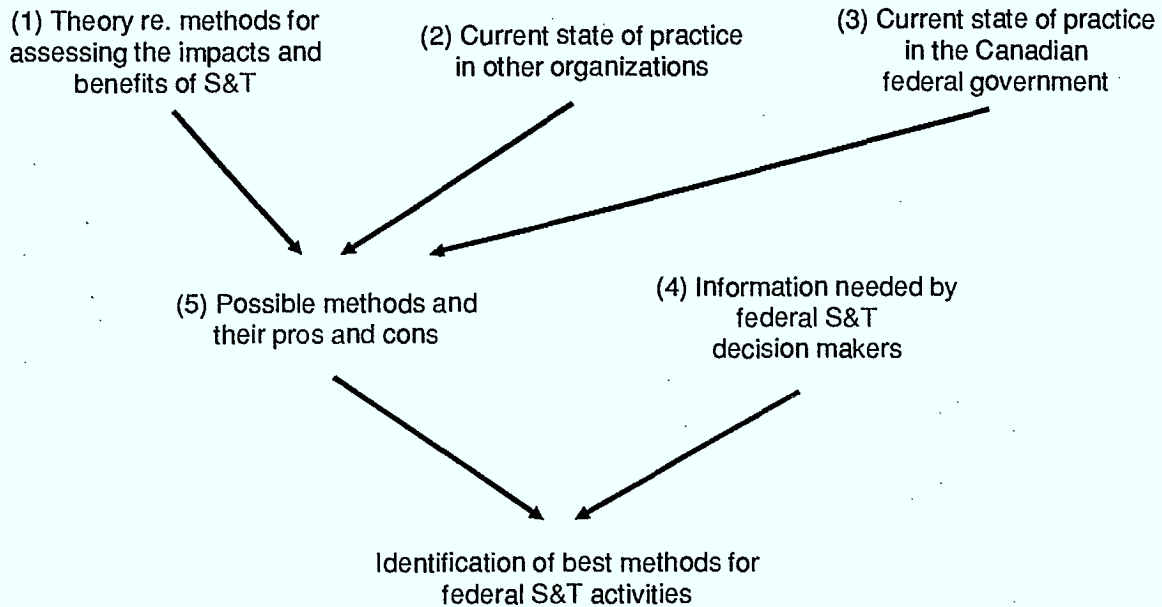
Further details regarding the main study activities are provided below.

Literature Review. This activity (items 1 and 2 above) involved an extensive review of both the published and unpublished literature. For the review of the published literature the study team used the services of NRC's Canada Institute for Scientific and Technical Information (CISTI). Searches were conducted using a number of key word groupings such as R&D, impact, technical innovation, benefit. As well, titles and abstracts of several key authors active in this field (Link, Mansfield, Tassey and Terleckyj) and key journals (e.g., Research Policy) were examined for relevancy to this study. CISTI then provided copies of specific articles selected from the vast array turned up by the search.

For the review of unpublished (grey) literature the study team was assisted by John Irvine of Science Policy Research Consultants in the United Kingdom. Mr. Irvine supplied copies of a large number of articles (some published but many unpublished) from his extensive library and summarized the highlights of these articles for the study team.

A total of approximately 100 articles were reviewed. The results of the literature review are summarized in Appendix A. That appendix contains a bibliography which lists all the articles that were found useful.

Review of Canadian Government Practices. The review of the current impact assessment practices and information needs of the Canadian federal government (items 3 and 4 above) focused on the practices of the major S&T departments and agencies. Information was collected using the following methods:

- All supporting departments and agencies were invited to contribute examples of recent work which included assessment of program impacts. Many did so. The study team reviewed approximately 30 documents illustrating the impact assessment practices of these departments. Most of these were program evaluation studies.

- Workshops were held with six departments and agencies chosen to broadly represent the major types of R&D activities carried out in the federal government (funding and performance of R&D activities, research training). These included:

  - Agriculture Canada;
  - Communications Canada (referred to as DOC in the remainder of the report);
  - The Canada Centre for Mineral and Energy Technology (CANMET);
  - Environment Canada (DOE);
  - The Natural Sciences and Engineering Research Council (NSERC);
  - The National Research Council (NRC).

For simplicity these are collectively referred to as the "participating departments" in the remainder of the report.

The purpose of these workshops was to discuss the impact assessment practices and information needs of these departments. The workshops generally included representatives of each of the following groups in the department:

  - senior managers for the department's main R&D program areas:

  - representatives of the department's program evaluation group (and representatives of internal evaluation groups within the main R&D program areas); and

  - senior policy people and decision makers in the department.

The information collected through this process is summarized in Appendix B. It should be emphasized that *this was not a comprehensive review of the impact assessment practices of the participating departments.* The scope of the study limited the data collection to the activities described above. This means, for example, that when we say in later sections "Department A used method B in study C" it should not be concluded that study C was the only study in the history of the department in which method B was used. Similarly, the information reported on departmental information needs should be regarded as indicative, not definitive.

Steering Committee. This study was overseen and directed by an interdepartmental Steering Committee, the membership of which is shown in Exhibit 1.1. The study team would like to express its appreciation to this committee for their constructive advice and guidance. Particular thanks are due to George Teather for his many helpful suggestions and assistance throughout the study.

1.3   Study Focus on R&D Activities

The original focus of this study, as outlined in the RFP, was on R&D activities, not the broader scope of S&T activities. Most of the data collection dealt with R&D activities, and, therefore, in the remainder of this report we refer to R&D rather than S&T. As discussed in Section 4.0, however, the methods described in the report can be applied to many non-R&D S&T activities.

Exhibit 1.1:   Steering Committee Members

William Smith (Chairman)
Corporate Policy and Evaluation
National Research Council of Canada

Bill Graham
Communications Development and Planning
Communications Canada

Nola Breithaupt
Science, Economic and Regional Development
Office of the Comptroller General

Carmen Foglietta
Research Branch
Agriculture Canada

Ann Cooper
Evaluation and Audit Branch
Department of Energy, Mines and Resources

Malcolm Drury
Coordination and Planning Division
Geological Survey of Canada

Greg Frenet
Science and Technology Strategy
Industry, Science and Technology Canada

Gilles Gauthier
Office of the Science Advisor
Environment Canada

George Teather (Secretary)
Evaluation and Review
National Research Council of Canada

2.0 Defining Impact Analysis

2.1 Definition of Economic and Social Impacts

An R&D activity is considered to result in an *economic impact* if the activity contributes to a change in income (for someone). The most common kinds of economic impacts resulting from R&D activities are:

- increased sales revenue, generally as a result of new or improved products, systems, or services; and

- cost savings.

An R&D activity is considered to result in a *social impact* if the activity contributes to a change in well-being (for someone). The most common kinds of social impacts resulting from R&D activities are:

- environmental enhancement;
- reduced health and safety risks;
- improvements in quality of life; and
- improved quality of and access to information.

Government R&D most commonly results in social impacts when the R&D is carried out in support of one or more government policy goals, such as protection and improvement of the environment.

Some economists would argue that all the impacts described above are economic impacts—the only problem is that it is difficult to quantify the "social impacts" in economic terms. In this report it is useful to maintain the distinction, however, because the impact assessment techniques that are best suited for R&D intended to lead to direct economic impacts are often different from the assessment techniques for R&D intended to lead to social impacts.

Note that, while the examples listed above are phrased as positive impacts, R&D can also result in negative impacts. For example, some people would view weapons development based on military research as a negative impact. Also, there are many cases in which R&D results in positive impacts for some people and negative impacts for others—for example, the development of a new product may be very beneficial for one company but may drive another company out of business.

2.2 Background to the Discussion of Impact Assessment Methods

The literature is filled with discussions of the difficulties of assessing the economic and social impacts of R&D, particularly R&D carried out in the public sector. These are summarized in a recent article as follows:[1]

- Public R&D is often undertaken for non-economic reasons, to achieve non-economic aims.

- R&D often produces benefits which take the form of "public goods" — that is, goods which are not bought and sold in the marketplace, such as clean air, public order, defense, and so on. Because prices do not exist for public goods it is difficult to measure their economic value.

- The methods of measuring impacts are particularly weak when it comes to public sector activities, especially when impacts take the form of qualitative improvements.

- R&D produces impacts by generating technological innovations. However, the connections between R&D and innovations are often long term, indirect, and unpredictable. Studies of technological innovations have shown them to often depend on research results that are decades old and, in many cases, in seemingly unrelated fields.

This last point is worthy of elaboration. Most of the impact assessment techniques developed to date have been built on the assumption that the R&D process can be represented as a linear model in which knowledge developed at one level of R&D moves forward in a straight line (and with a reasonably fixed time lag) to more advanced R&D activities—i.e., basic research flows into applied R&D activities and then into the development of products or processes. However, with the exception of R&D in specific, focused areas, this model has proven to be an inadequate way of conceptualizing the vast majority of R&D activities and, consequently, of identifying and measuring their economic and social impacts.[2] A number of recent studies suggests that R&D (especially, less applied R&D) may incorporate forward knowledge diffusion links, lateral diffusion links (i.e., research in one field may affect research in another), and/or backward diffusion links (e.g., product and process development may lead to modifications in basic and strategic research). Furthermore, the knowledge diffusion and innovation process tends to exhibit significantly non-linear characteristics and is often discontinuous. Existing impact assessment models are constrained in their ability to account for the various diffusion links and the number and types of paths along which the benefits associated with R&D may arise.

---

[1] *Economic Returns to R&D: Methods, Results, and Challenges*, Keith Smith, Innovation Studies Group, Science Policy Support Group, Oslo, Norway, November, 1991.

[2] *Innovation Is Not A Linear Process*, Steven Kline, **Research Management**, Volume 28, #4, August, 1985, Pages 36-45.

To complicate this situation, the literature also notes that many activities of government laboratories and R&D programs contribute to socioeconomic benefits in more subtle ways than direct technology transfer. For example, benefits can result from helping to bridge gaps in the innovation chain, providing access to instrumentation, providing advice, informal cooperation, and so on.[3]

A further complication is the fact that in order for R&D to produce impacts not only does the R&D have to lead to potentially useful results but many other things have to happen as well. Most of these are totally beyond the control of R&D managers—e.g., availability of risk capital, favourable market conditions, and so on. The illustration on the following page, adapted from Tassey,[4] shows the activities that have to occur in order for R&D (large box) to lead to economic benefits (value added). (This illustration also shows in the circles the typical roles of government in supporting this process.) This vastly complicates the situation for those trying to predict the impacts of on-going R&D or R&D that might be carried out in the future. Even for R&D carried out in the past this means that it is often difficult to identify the role of the R&D in contributing to the socioeconomic impacts which ultimately occur.

While the difficulties discussed in the literature and summarized above tend to paint a rather gloomy picture of the potential for **comprehensively** assessing the social and economic impacts of R&D, the literature also makes it clear that a number of practical and credible assessment techniques are available. While these techniques may not be able to capture **all** of the impacts associated with an R&D activity for the reasons given above, they can at least provide guidance in qualitatively identifying potential impacts and, in many cases, can provide reasonably accurate quantitative estimates. As will be seen, this is born out by the experience of other countries, as well as the Canadian experience. What needs to be kept in mind, then, is that these methods may tend to underestimate the benefits of R&D (unless, of course, there is something inherent in the method that skews the results in the direction of overestimating benefits).

The assessment methods which are most appropriate in any particular circumstance obviously depend on the type of R&D being assessed, as well as other factors. As a result of the literature review, the following were identified as being the main factors that determine which methods are most appropriate:

(1)   The time frame for the assessment—specifically, whether the social and economic impacts of an R&D activity are being assessed prior to or after the R&D has been completed.

(2)   The type of R&D—basic and strategic research, applied research, or product/process development.

---

[3]   For a discussion of these kinds of benefits see *Federal Laboratories and Competitiveness: An Evaluation Model*, Nola Breithaupt, Office of the Controller General of Canada, May, 1992.

[4]   Tassey, G., *Technology Infrastructure and Competitive Position*, Norwell, Mass., Kluwer Academic Publishers, 1992, page 261. We have added the circle on the lower right. .

(3)   The purpose of the R&D—in particular, which of the following categories
        best describes the R&D activity:

Category 1-R&D Infrastructure: R&D undertaken to contribute to society's R&D
infrastructure (through the development of research equipment, the maintenance of
research capability, and so on).

Category 2-Policy Development: R&D undertaken to provide information needed to
develop government policies or standards and regulations.

Category 3-Policy Attainment: R&D undertaken to provide information, products,
processes, and systems that will contribute to the attainment of government policies (such
as environmental protection, economic development, and so on).

Category 4-Industrial Development: R&D undertaken in support of industrial
innovation.

The main types and purposes of the R&D carried out by the participating departments are shown
in Exhibit 2.2.

In general, the potential for finding methods for satisfactorily assessing economic and social
impacts is greater for R&D activities which have been completed, for activities toward the more
applied end of the spectrum, and for the higher numbered purposes of the R&D (e.g., Category
4 rather than Category 1).

R&D Type	R&D Purpose			
	Category 1 R&D Infrastructure	Category 2 Policy Development	Category 3 Policy Attainment	Category 4 Industrial Development
Basic/Strategic	AC, DOC, CANMET, DOE, NSERC, NRC	AC, DOC, CANMET, DOE, NRC	AC, DOC, CANMET, DOE, NSERC, NRC	AC, DOC, CANMET, NSERC, NRC
Applied	AC, CANMET, DOE, NSERC, NRC	AC, DOC, CANMET, DOE, NRC	AC, DOC, CANMET, DOE, NSERC, NRC	AC, DOC, CANMET, NSERC, NRC
Development	AC, CANMET, DOE, NRC	AC, DOC, CANMET, DOE, NRC	AC, DOC, CANMET, DOE, NRC	AC, DOC, CANMET, NRC

AC	=	Agriculture Canada
CANMET	=	Canada Centre for Mineral and Energy Technology
DOE	=	Environment Canada
NSERC	=	Natural Sciences and Engineering Research Council
NRC	=	National Research Council
DOC	=	Communications Canada

3.0 Impact Assessment Methods

The methods which have been found to be the most useful for assessing the impacts of R&D are discussed in Sections 3.1-3.6. Some other methods which may be useful in certain situations are discussed in Section 3.7. Methods that were reviewed and found to be not particularly useful are discussed in Appendix A.

3.1 Modified Peer Review

3.1.1 Description of the Method.

Traditional peer review involves the assessment of the quality (scientific merit) of the research by scientific experts (peers) in the specific research field. While peer review is the most widely used method of evaluating R&D, it does not directly involve the assessment of its economic and social impacts.

Modified peer review involves the joining of some form of socioeconomic impact assessment with traditional peer review. The simplest type of modified peer review involves the selection of peers who can comment not only on the scientific merit of the research but also on its economic and social impacts. The peers are then asked to address both issues.

It is often not possible to find scientific peers who have the expertise to provide valid assessments of economic and social impacts. In these cases the standard practice is to modify the normal peer review process by including non-peer members in the expert group, such as economists, social scientists, and civil servants. For example, a recent modified peer review of the Swedish salix (fast growing willow) R&D program involved the assessment of the scientific merit of the research by university researchers from outside Sweden combined with the assessment of the potential economic impacts of the research by a well-known Swedish agricultural economist.

The most common structure of a modified peer review involving scientific peers and socioeconomic experts is that the two reviews are carried out independently and the results are then brought together in an integrated report by a third party. Sometimes, however, the assessments of the scientific peers are provided to the socioeconomic experts, who then take these opinions into consideration in carrying out their own review.

Modified peer reviews generally do not go beyond a basic question and answer format, often facilitated by interview techniques. However, if more than one socioeconomic expert is involved in commenting on a particular R&D activity, the method can involve an iterative discussion process which attempts to reach consensus regarding the social and economic benefits (possibly also including such items as probabilities and occurrence times).

While useful for at least the qualitative assessment of benefits, modified peer reviews have some major weaknesses. The main one is that generally a relatively small number of individuals are involved, and, as a result, it is difficult to gather a sufficiently broad base of knowledge to credibly comment on the economic and social impacts of the R&D. The problem is exacerbated if the R&D activity being assessed has a broad range of possible uses and users. This problem can be effectively overcome by broadening the modified peer review through the application of an extended set of interviews and questionnaires, as discussed in Section 3.2.

3.1.2 Applicability

This method is equally well suited for assessing R&D carried out in the past, on-going R&D, and R&D that is being considered for the future.

The method can be used for basic and strategic R&D as well as applied R&D and development. It can be used for assessing R&D carried out for each of the four purposes described in Section 2.2., although it is least well suited for Category 1 R&D, in which the intended users and uses of the R&D may be less well defined. It is most useful for assessing R&D projects and programs, and not particularly useful for assessing larger R&D activities (e.g., the nation's total research effort in agriculture).

3.1.3 Examples of Use

Use in Other Countries. Modified peer reviews are widely used in European countries, often in combination with some other method. Generally these reviews are conducted in accordance with the process described above. However, some notable exceptions exist. For example, the Swedish Natural Science Research Council uses a two stage modified peer review panel, composed of foreign and domestic experts, to assess both scientific/technical outcomes and commercial applicability. This same method is used to a lesser extent in Finland, Norway, and Denmark for the assessment of some basic research fields. The Dutch Organization for the Advancement of Pure Research, in arriving at its funding decisions, uses a two-stage version of the modified peer review process which employs the Delphi technique to determine the probability of obtaining benefits from the research activities under consideration.[1]

Use in the Canadian Government. This technique has not been heavily used for assessing past R&D by the participating departments. DOC and DOE have each used it in its simplest form (one reviewer commenting on both scientific quality and usefulness/relevance) in program evaluations, and NSERC uses it in the assessment by grant review panels of the final reports for research projects supported by their Strategic Grants program.

---

[1]   Luukkonen-Gronow, T., *Scientific Research Evaluation: A Review of Methods and Various Contexts of their Application, R&D Management*, Volume 17, Number 3, 1987.

It is more heavily used for assessing on-going and future R&D. Most of the participating departments—in particular, Agriculture Canada, DOC, CANMET, and DOE—have advisory committees consisting of representatives of the primary intended users of the research results who comment on the relevance and usefulness of on-going and potential future R&D. This could be considered a type of modified peer review exercise, although generally these committees do not explicitly consider the scientific quality of the research.

There have also been some ad-hoc modified peer reviews of on-going research in some of the participating departments. For example, in Agriculture Canada, there was a recent review of the food safety and toxicant research at the Central Experimental Farm carried out by a review panel which included both academic and industry representation and which considered both the scientific quality and the potential usefulness of the research.

The heaviest users of this method among the participating departments appear to be NSERC and NRC. For NSERC's targeted programs, including the Strategic Grants program and the Cooperative Research and Development (CRD) program, there are expert panels which assess both the scientific merit and the potential user relevance of proposed research projects. For the CRD program, this review includes an analysis of the potential for Canadian companies to capture the benefits of the research where appropriate. (For the more developmental projects supported by this program, the program actually requires a market study, but only about 5% of the projects fall into this category.)

At NRC modified peer review serves as the foundation of the normal program review process. In this process peer reviews are carried out in parallel with program evaluations, and the results are brought together at the end of the review exercise. They have found this approach very useful.[2]

3.2 User Surveys

3.2.1 Description of the Method

The most common method for assessing the economic and social benefits of R&D involves the use of surveys (through either interviews or questionnaires) of the primary intended users of the research results. There are two basic types of such surveys:

- Surveys in which the respondents are involved with the research organizations in research collaboration or some other form of active interaction (e.g., the active exchange of information). These surveys are sometimes called "client surveys".

---

[2] *NRC's Approach to the Evaluation of R&D Programs*, paper presented by William Smith at the NRC Seminar on the Evaluation of R&D Programs, December 14, 1989.

- Surveys in which the respondents are selected to be a representative sample of the broader group of primary intended users of the research results. (This group is sometimes called "beneficiaries".) These surveys are sometimes structured in such a way that the direct clients (as defined in the first point) are a certain percentage of the total sample of primary intended users.

In the second type of user survey it is sometimes necessary to use experts rather than users to review some of the research—for example, in the case of research for which the primary intended users are not sufficiently familiar with the research to be able to answer the questions posed.

This method has the advantage of providing a more systematic review by using standardized interviews and questionnaires and gathering the views of a wider number of people, thereby overcoming the restrictions involved in using a modified peer review procedure with a limited number of participants. The technique also provides the advantage that quantitative indices can be formed if the questions are amenable to scoring, thereby providing a convenient method of making comparisons among R&D projects.

The main issues that have to be dealt with when using this method are:

- which type of user survey to implement;
- how to identify the users;
- how to structure the survey sample;
- when to use experts instead of users;
- how to ensure the validity of the results; and
- how to interpret the results.

The issue of the validity of the results is particularly tricky. When users (either actual or intended) of R&D programs are asked to comment on the relevance and usefulness of the R&D, they sometimes have a tendency to be more positive than is born out by the facts. This is generally referred to as the problem of "grateful testimony".[3] There are a number of techniques that can be used for dealing with this problem, which are discussed in Appendix A. It is noteworthy that these techniques have only been used on a widespread basis within the past several years.

The question of how to interpret the results can also be tricky. Because this method is based on surveys, it is generally not possible to ask detailed questions regarding economic and social impacts of the R&D. Therefore, one obtains results such as: 50% of the intended users for a particular group of R&D projects rated the usefulness of the projects as "high"; 30% rated the usefulness as "medium"; and 20% rated the usefulness as "low". What does this mean? Is this

---

[3] *Relevance and Limitations of Various Methods and Approaches to R&D Evaluation*, Bennett, D and I. Jaswal, paper presented to the NRC Seminar on the Evaluation of R&D Programs, December 14, 1989.

a good result or a bad result? Again, there are techniques for dealing with this (see Appendix A), but these techniques have only been applied relatively recently.

While the use of this method has proven valuable for gathering information on the perceptions of the intended users regarding the value of R&D activities, there are two cautions regarding its use which should be noted. First, in developing questionnaires and interview guides one has to be careful to standardize the questions in a manner which facilitates analysis but does not place undue constraints on the amount of information that can be obtained, thereby resulting in the collection of trivial information. Secondly, the individuals involved in developing the interview guides or questionnaires must have a clear understanding of the nature of the R&D activity being assessed. If not, the potential exists that the wrong questions may be developed and, again, trivial or misleading results could be obtained.

3.2.2 Applicability

This method is most useful for past R&D and on-going R&D, for which the respondents can be given concrete descriptions of the research projects and, for past R&D descriptions of the results. (For on-going R&D respondents can often be given descriptions of some initial outputs, such as a list of publications to date.) For research which is only being considered for the future the respondents (who are intended users but not necessarily experts) may not have a sufficiently good idea of what types of research results may be obtained to be able to comment on the likely impacts.

The method is useful for assessing the impacts associated with R&D activities near the applied and development end of the R&D spectrum. For basic and strategic research activities information concerning the potential uses of the research may be too vague and too limited for users to be able to assess the potential impacts. Also, the method is best suited for the assessment of R&D projects and programs and for Category 2 to Category 4 R&D.

3.2.3 Examples of Use

Use in Other Countries. The majority of the countries whose assessment practices were reviewed employed user surveys of some type, often combined with some other method. The most common combination involved user surveys combined with case studies (discussed below) to gather more detailed information on the consequences of particular R&D activities.

Use in the Canadian Government. User surveys have been heavily used in evaluations of Canadian government R&D programs and other R&D assessments. The experience of the participating departments with this method can be summarized as follows:

- Agriculture - This method was used in three recent program evaluations (although in two of these the focus was more on program rationale issues than program impact issues) and is currently being used for both impact assessment and planning purposes in a study for the Central Experimental Farm.

- DOC - The recent evaluation of the DOC research labs included two user surveys—one of direct clients and one which included non-direct client "beneficiaries". As well, DOC is surveying clients and intended users as to their needs in the major program review which is in progress.

- CANMET/GSC - This technique has been used in all recent program evaluations (CANMET, GSC, PERD).

- DOE - The technique was used in two program evaluations. (It has also been used in this Department for the assessment of some non-R&D activities—for example, in client satisfaction studies undertaken by the Canadian Parks Service and the Canadian Wildlife Service.)

- NSERC - The recent interim evaluation of the Networks of Centres of Excellence program included a survey of the industrial participants in the networks. NSERC is planning to implement regular user surveys for projects that have been supported by its Research Partnerships program.

- NRC - User surveys have been one of the main data collection methods for almost all program evaluations (e.g., Division of Biological Sciences, IRAP, Engineering Sector). These have been surveys of direct clients.

These user surveys have generally involved satisfaction ratings, ratings of the importance of the research to the client organization, information on impacts of the research on their organization (e.g., degree of technology transfer, impact on sales), and information on the implications for them if the government were to carry out either less (including zero) or more research in this area.

Five of the six participating departments have found the information provided by this method to be useful. The DOC labs did not find the information particularly useful because the study was overtaken by other events, such as the high profile Lortie study of government S&T activities performed for the Prime Minister's National Advisory Board on S&T (and also because the results were relatively positive and no major changes were recommended).

3.3   Benefit-Cost Methods

3.3.1 Description of the Method.

Benefit-cost analysis provides a strong theoretical framework for analyzing the economic and social impacts of R&D activities. It is always carried out on a project by project basis, and it seeks to assess the project in terms of both the economic and social benefits generated for society, as well as the economic and social costs incurred by society to execute the project. The utility of the project from society's point of view is then expressed in terms of the net benefits (i.e., gross economic and social benefits minus economic and social costs) generated by the project.

There are several variants of benefit-cost analysis—including rate of return calculations and net present value calculations—but these all involve essentially the same techniques.

In defining the costs and benefits employed in this type of analysis, very specific definitions and methods of evaluation are used. A few of the key points are:

- There are three main types of costs associated with an R&D project that must be taken into consideration in the analysis: the costs of generating the research results; the costs of introducing and supplying the results to end users; and the costs incurred by the end users to implement the results. Each of these cost categories needs to be identified and included in the analysis.

- The benefits of the R&D project, which result from the new or improved products, processes, or systems which result from the research, are valued at the price society is willing to pay for them. The assessment of benefits includes not only those for which prices are paid, but also benefits associated with increased educational and training opportunities, reduced environmental damage, improvements in health and safety, and so on, even though in many instances it may not be possible to associate an explicit value with such benefits.

- In addition to the definition and evaluation of the costs and benefits associated with an R&D project, the probabilities associated with the realization of these costs and benefits must be determined, and the time sequence of the relevant costs and benefits must also be determined. These items are necessary, since the expected value stream of annual costs and benefits must be discounted to their present values in order for the calculation of net benefits to be carried out.

The requirements for successfully carrying out benefit-cost analysis of R&D activities are very demanding. The two that are the most demanding are:

- It must be possible to associate probabilities with research outcomes and subsequent applications; and it must be possible to determine the time sequence over which these outcomes and their applications will be realized.

- The calculations must be based on **incremental** benefits and costs—i.e., only those which would not have occurred in the absence of the R&D project. It is often very difficult to meet this requirement, because many social and economic benefits result from a combination of complimentary R&D investments, incurred over substantial periods of time, and it is often not possible to isolate the influence of a specific R&D project.

Given the inherent uncertainty of R&D, many evaluators find that benefit-cost analysis is impractical and far too technically demanding for most R&D projects, except possibly those near

the applied and product/process development end of the R&D spectrum. The main value of benefit-cost analysis in R&D impact assessment is probably that it offers a systematic framework for identifying the costs, benefits, and wider implications of R&D and its exploitation. It is often used to illustrate the kinds of benefits that can occur from R&D.

In the past few years there have been an increasing number of "partial benefit-cost analyses" of R&D programs carried out. In this type of analysis, one selects R&D projects which appear to have been (or have the potential to be) "big winners" economically. One then carries out an individual benefit-cost analysis for each of these projects and calculates the sum of the net benefits for the selected projects. If the net benefits from these projects is large—for example, if it is large enough to cover the R&D costs associated with the remaining projects in the program—one may be able to reasonably conclude that the benefits associated with the program's R&D outweigh the costs.

3.3.2 Applicability

Benefit-cost methods are much more appropriate for past research than for on-going or future research. In order to apply benefit-cost analysis to on-going or future research, one must have a good idea of the likely outcomes of the R&D, their probabilities of occurrence, when they will occur, how and when they will be applied, and the market for the products or processes developed on the basis of these results. Obviously such information is quite speculative for most R&D which is still underway or which is only being planned for the future. The method can be used, however, for on-going or future R&D within those sectors where the connection between the R&D and sectoral economic impacts are clearer and more direct, such as agriculture.

As noted above, benefit-cost analysis is best suited for R&D projects near the applied research and development end of the R&D spectrum. It is most suitable for Category 4 R&D activities (and in some instances Category 2 activities if, for example, standards or regulations that are developed apply to a fairly narrow industry group and Category 3 R&D if the links between the R&D and the attainment of the policy can be identified and the benefits of the attainment of the policy can be quantified). Finally, since it is carried out on a project by project basis and the benefit-cost analysis for even one project can be fairly time consuming, it is suitable only for individual projects or relatively small collections of individual projects (i.e., small programs).

3.3.3 Examples of Use

Use in Other Countries. It would appear that the only two countries which have made a significant amount of use of benefit-cost methods for R&D impact assessment are Australia and Canada. It is rarely used in other countries, and when it is, it is generally used on an ad hoc

basis for evaluating the short-term development-oriented aspects of R&D activities.[4] The use of the method in Australia[5] is very similar to its use in Canada (discussed below).

Use in the Canadian Government. Canada is one of the world leaders in the application of benefit-cost analysis to assess past R&D activities. Successful partial benefit-cost analyses (see above for definition) have now been carried out for the NSERC Strategic Grants program, the DOC research laboratories, and the CANMET laboratories. Agriculture Canada funded several studies in the mid-1980s which used benefit-cost techniques to evaluate the benefits of past agricultural R&D.[6] CANMET has also used benefit-cost analysis successfully in a number of program evaluations studies, including the evaluation of the Industry Energy Research and Development program (which included benefit-cost analyses of 37 IERD projects) and the evaluation of several demonstration programs. Even when this method is not used, the feasibility of using it is often investigated. For example, in the PERD evaluation the feasibility of a partial benefit-cost analysis was investigated, but it was decided not to proceed because of the difficulty of attributing the benefits of PERD-supported R&D projects (which are also generally supported by funding from some other source) to the PERD program.

The clients for all of the partial benefit-cost studies noted above found them to be very useful, particularly for documenting in a rigorous way the kinds of benefits that can result from R&D activities. The Evaluation Branch of EMR found the benefit-cost work done in the evaluations mentioned above to be very useful, but they cautioned that this method would be less useful for assessing R&D projects that are further away from commercial implementation, because of the inherent uncertainty of the assumptions that have to be made.

An example of the use of benefit-cost methods for prospective assessment was found in Agriculture Canada. The Research Branch is currently developing a return on investment model to calculate the expected value of economically-oriented R&D projects. This model will be used mainly for resource allocation purposes. It will incorporate consideration of the probability of success of the R&D, the potential economic impact, and the cost. (The model is only one tool that is being planned to assist with project and project area selection. There are also some non-quantitative guidelines for project selection.)

---

[4] *Evaluation of Research and Development: Current Practice and Guidelines*, W. Krull, D. Sensi, D. Sotiriou, Commission of the European Communities.

[5] For a typical example see *Rural Research - The Pay-Off*, CSIRO Occasional Paper No. 7, May 1992. This report summarizes the results of benefit-cost studies of 10 "winner projects". (It also contains a good discussion of methodological and practical considerations). For an example of the use of benefit-cost analysis for prospective impact assessment see *Economic Evaluation of CSIRO Industrial Research*, Australia Bureau of Industry Economics, 1992.

[6] These are summarized in the report *Economic Evaluation of Agricultural Research in Canada*, Research Branch, Agriculture Canada, 1990.

The Department recognizes the inherent inaccuracy of this method as a result of the need to make a number of critical assumptions about things which cannot be predicted with a high degree of accuracy. However, it is felt that a major benefit of the technique is that it forces the right questions to be raised at the time projects and project areas are selected.

There are, of course, a number of examples of the use of benefit-cost methods for the prospective assessment of the impacts of major government investments which are R&D oriented, such as the space program or the KAON project.[7] We are not aware of any such examples in the participating departments, however.

3.4   Case Studies and Histories

3.4.1 Description of the Methods.

Case studies represent one of the most useful methods for examining the relationship between R&D and its associated economic or social impacts. Case studies involve a detailed and thorough analysis of particular R&D projects or programs and seek to track and document the evolution of economic and social impacts associated with these activities. They are generally conducted in conjunction with other methods—e.g., benefit-cost assessments of case study projects—or as a method of validating or illustrating results from interviews and surveys. The advantage of case studies is that, when they are carried out in sufficient number and in sufficient detail, they represent probably the best chance of fully identifying the relationship between R&D activities and the resulting economic and social impacts. The problem with this method from the point of view of the assessment of research is that, since case studies relate to specific R&D projects or a specific sample of projects, it is difficult (generally impossible) to aggregate the "results" from a group of case studies or to generalize the results to larger R&D activities, such as R&D programs or major research efforts.

Histories are a variant of the case study approach. They involve attempting to document the impacts of R&D from both directions—i.e., tracing out historically the key events from research to application to development and vice versa taking a particular product or process and looking "backward" to try to identify the key R&D results that contributed to the development. The best known history-type methods are the Hindsight Method and the Technology in Retrospect and Critical Events in Science (TRACES) Methods.[8] These methods were originally developed in the 1960's and have recently been used by the Office of Naval Research in the United States and the Science and Engineering Research Council in the United Kingdom. Although they can provide good insights into the impacts of R&D, there are a number of problems associated with their use. The main criticisms are that they often falsely assume causal connections between

---

[7]   See, for example *KAON Economic Assessment*, The DPA Group (now the ARA Group), February, 1990.

[8]   See *Research Impact Assessment*, Office of Naval Research, 1992, for a discussion of these methods.

---

events, incorrectly assign equal weights to events, and ignore historical dead ends. In addition, the results of these analysis tend to be sensitive to the timeframe that is selected.

3.4.2 Applicability

By their very nature case studies deal with R&D that has been carried out in the past, since they involve documenting the connections between the conduct of the R&D and its socioeconomic impacts.

Case studies are usually applied in assessments of Category 2 to Category 4 projects. They are usually—but not exclusively—used for projects toward the applied end of the R&D spectrum. Case studies are often the best method for assessing the impact of government R&D activities aimed at policy development and the development of standards and regulations (Category 2 R&D).

The history method is used for the same categories of assessments as case studies. Because of the focus of the history method on identifying key events which link R&D with development, the method tends to focus more on process. (If nothing else, these methods have been useful in highlighting the complexity of the problems faced in linking R&D to economic and social impacts.)

3.4.3 Examples of Use

Use in Other Countries. There is very little mention of case studies in the literature, possibly because case studies are not very exciting methodologically—you basically go out and get all the relevant information you can and write it up. On the other hand, in our review of actual R&D assessment studies in other countries, we encountered many examples of the use of case studies.

Histories, in the other hand, have not been widely used in other countries, the main exceptions being several major exercises in the United States associated with military and aerospace R&D.

Use in the Canadian Government. Case studies have been fairly heavily used in assessments of R&D in three of the participating departments but rarely used in the other three. The three that have used case studies are:

- CANMET/GSC - Used in the evaluation of the GSC, PERD (24 case studies), and CANMET (essentially 60 case studies). In addition, the impact assessment system planned by the PERD program for completed projects includes for each year:

    - the preparation of case studies of major environmental benefits due to PERD projects which are realized in that year; and

    - the preparation of case studies of the most important government uses of PERD results in that year.

- NSERC - The evaluation of the Research Partnerships program was based primarily on case studies (44). NSERC also prepares case studies of a number of research projects for various promotional documents (see Appendix B).

- NRC - Used in several evaluations (e.g., Division of Biological Sciences, Biotechnology Program). Case studies have also been done outside of program evaluations for strategic planning purposes.

Case studies have generally been found to be useful for program evaluation purposes. This was particularly the case for the PERD evaluation, the NSERC Research Partnerships evaluation, and the NRC Biotechnology Program evaluation. They have clearly been useful to NSERC for promotional purposes. Also, NRC officials reported that they have been useful for strategic planning purposes—the identification and documentation of high impact projects has provided useful information for identifying future projects and project areas.

Histories have not been widely used in Canada, although there have been a few exceptions. For example, in the PERD evaluation a history approach was used to try and document the linkages between energy R&D and increased security of energy supply and also the relationship between energy R&D and increased environmental quality associated with energy exploitation and use. These histories were interesting although not wildly successful in identifying the role of R&D in the eventual outcome.

3.5   Partial Indicators of Impacts

3.5.1   Description of the Method.

This method involves the collection of information (generally relatively readily available) for a number of items each of which provides some insight into the extent of the socioeconomic impacts resulting from the R&D. For example, for an R&D program the information that is collected can be information on inputs (program funding, number of people involved in running the program, and so on), program activities, program outputs, or program impacts themselves. In effect, one sets up an information collection system for the program, and, once the information has been collected, it is organized and presented in a way that enables people who are reviewing the information to draw conclusions regarding the impacts of the program (and especially changes in the impacts over time).

For example, the Swedish Plastics and Rubber Institute has set up a system of partial indicators which includes the following indicators for assessing the impacts of past R&D:

- *The percentage of projects completed during the past year for which the technical goals were met (or exceeded).*
- *The number and percentage of available project reports sold to industry during the past year.*
- *The number of projects that have had a documented impact on industry during the past year.*

- *The percentage of projects completed within the past year for which it is highly likely that the results will ultimately be used by industry* (this one requires an external assessment).
- *The number of patent applications during the past year.*
- *The number of licence contracts signed during the past year.*
- *The amount of revenue obtained during the past year from licences.*
- *The number of products developed during the past five years that are being further developed or marketed by industry.*

The advantage of partial indicators is that the information required to specify the indicators is relatively easy to collect. Their disadvantage is that they only provide a very partial picture, and while this can be useful for program monitoring purposes, it is generally not sufficient for demonstrating the impacts of the R&D—or even understanding what they have been.

3.5.2 Applicability

Systems of partial indicators are an appropriate method for assessing past, on-going, and future R&D. They are implicitly used by many research managers in selecting projects and project areas (i.e., indicators such as: Is there an identified need? Is there an identified client?), and they are heavily used in the monitoring of on-going R&D projects and programs.

This method is applicable to all types and all purposes of R&D, and it is probably the best method for more fundamental R&D and for R&D oriented toward the development and maintenance of research infrastructure. It can easily be combined with other methods—for example, if a modified peer review is carried out, the opinions of the reviewers regarding the likely usefulness of the research can be incorporated in the indicator system.

3.5.3 Examples of Use

Use in Other Countries. Partial indicators are widely used in other countries, but generally only by research program managers for program monitoring purposes, and more often than not the system of indicators is fairly informal (and rarely documented). The indicators are rarely used for reporting purposes (e.g., to central agencies). One example of a case in which partial indicators are used for both purposes is the formal performance indicator system recently set up by the Swedish National Board for Technical Development. (The agency was directed to set up such a system by the Swedish Parliament.)

Use in the Canadian Government. As in other countries, partial indicators are generally used by research managers on an informal basis for monitoring purposes. The following examples of the more formal use (or potential use) of partial indicators were identified in this study:

- Agriculture—A formal performance indicator system was developed in 1987 for the St. Hyacinthe Food Research Centre. This system includes a number of indicators of potential use and impact for the on-going R&D: number of joint projects with industry, number of projects focused on marketable applications, degree of

involvement of industry in managing the research program, number of projects for which there was a clear definition of the economic impacts anticipated before project approval, and so on.

- CANMET/GSC—CANMET has a system of performance indicators intended to measure its effectiveness in working with industry. These are, therefore, at least partial indicators of potential economic impacts. The GSC uses similar performance indicators. In a recent GSC workshop on performance measurement senior managers concluded that a performance measurement process based primarily on partial indicators best suits the organization's needs. The PERD program has accepted in principle a performance indicator system for on-going projects that includes the following partial indicators of potential impacts:

  - percent of PERD budget devoted to projects for which clients have been identified;

  - percent of PERD budget devoted to projects for which uses and impacts have been identified; and

  - for projects intended to have economic benefits—percentage of budget for these projects which is devoted to projects expected to have positive net benefits.

For the first two indicators there is a planned system of checking the extent to which likely users and uses have been accurately and realistically identified (see Appendix B).

- NRC—The Portfolio Analysis system recently implemented by NRC's Institute for Environmental Chemistry is an example of a performance indicator system for on-going projects. This system involves the review of all projects and the scoring of projects on each of four criteria (each of which is broken down further into individual sub-factors which are scored):

  - Strength of IEC capability;
  - Strength of methodology/approach;
  - Potential "worth" of successful project; and
  - Marketplace "success factors".

One of the sub-factors included in the third criterion deals with potential social impacts, and all of the sub-factors under the fourth criterion deal with potential economic impacts (see Appendix B). The principle aim of the system has been to develop a database that will help the Institute management to better understand their projects and their potential impacts.

- NSERC—In most NSERC evaluations ratings by researchers of the extent to which the results of their research have been used by industry or government have been obtained. Also, the Networks of Centres of Excellence program has a formal system of partial indicators for each network (currently being revised) which includes some indicators of impacts of the R&D—e.g., extent of industry participation, number of patents, number of technology licences.

It is also noteworthy that in several of the departmental workshops it was stated that it would be very useful to have the use of each past R&D project documented (or, as a minimum, whether or not there had been a known use of the results). This is, of course, the most important partial indicator for past R&D.

3.6   Integrated Partial Indicators

3.6.1 Description of the Method

This approach is also often called "weighted multiple criteria analysis" or "scoring analysis". It is most commonly used for the assessment of future R&D options, and it provides a way of incorporating and ranking a wide variety of factors that influence the options that are selected. Conceptually, this method differs little from a checklist of partial indicators of impacts. The difference between the two approaches lies in the fact that there is some sort of system for "adding up" the partial indicators and arriving at a "bottom line score" for each potential R&D project or project area under consideration. The most common approach is to evaluate each project with reference to a specific set of criteria/questions (the partial indicators). Each criterion is then assigned a numerical weight, which enables the array of R&D projects or project areas up for consideration to be ranked in order of priority according to the sum of the numerical values assigned to the various criteria.

This method has a number of attractive features. First, it forces R&D decision makers to determine the criteria for assessing what makes for a good R&D investment. Then, it forces the decision makers to evaluate, for each project or project area, all the significant factors which have a bearing on the "worth" of the R&D and to make conscious trade-offs among multiple goals. Finally, it compels decision makers to rank R&D projects in terms of their relative importance. In addition, this approach offers the advantage that it provides a quantitative decision rule for deciding on R&D proposals.

The difficulty with this approach involves the potential arbitrariness and subjectivity in assigning weights to the various criteria. The elicitation of these weights must be carefully structured to avoid bias. This problem is especially evident in instances where decision makers may conclude that the score given a particular project is not in accordance with "common sense". In this event, the decision maker may choose to manipulate the weights assigned until a result emerges which appears to be more reasonable. Such manipulation of weights is not necessarily inappropriate and may be done by a single decision maker or by group consensus (which is preferred). However, it is important to guard against manipulating the results until they merely reflect a priori assumptions and historical preferences. (An additional difficulty with this

approach is that it is not very well suited for ranking R&D projects in significantly diverse research areas.)

3.6.2 Applicability

This method is intended to be used for assessing R&D under consideration for the future. It could also be used for assessing on-going R&D if, for example, resource constraints made it necessary to discontinue work on certain projects or project areas. It is not really appropriate for assessing R&D carried out in the past.

The method works best for research toward the applied/development end of the scale. (The weights are generally too difficult to assign for basic/strategic research.) It can be used for Category 2 to Category 4 R&D, although it is better suited for comparing projects within categories rather than across categories.

3.6.3 Examples of Use

Use in Other Countries. Methods of integrated partial indicators are commonly used in other countries for priorizing potential research projects and project areas. Straightforward scoring systems are the most common. A somewhat more interesting example of an integrated partial indicator method is used by the Commonwealth Scientific and Industrial Research Organization (CSIRO) in Australia. This system is used for assigning relative priorities to broad research fields—e.g., research related to the minerals industry vs. research related to transport.

The method is fully described in the Australia Bureau of Industry Economics Document previously referenced. Briefly, each research field is ranked on two criteria, "feasibility" and "attractiveness". Feasibility is an R&D factor which is intended to incorporate consideration of the R&D potential (i.e., the technical potential of the relevant research fields) and CSIRO's R&D capacity (i.e., their capability to carry out the R&D and Australia's international research competitiveness in this area). Attractiveness is a socioeconomic factor which incorporates consideration of the potential benefits of the R&D (both commercial, such as size of market, and non-commercial, such as health and safety improvements) and also Australia's ability to capture these benefits (for example, the receptor capacity of Australian industry for commercially oriented research and the likely implementation of the relevant research by public sector bodies for non-commercial research). For each research field the feasibility rating and the attractiveness rating are determined through a consensus process involving the senior managers of all of CSIRO's research establishments. The highest priority is assigned to research fields which rank high on both feasibility and attractiveness.

Use in the Canadian Government. The only example of the use of an integrated partial indicator system that was identified among the participating departments was the "rationalization exercise" currently being carried out by DOC. This exercise involves looking at all the R&D projects being undertaken by the labs and rating them against a set of criteria. These include criteria related to potential social impacts (e.g., degree of support for departmental policies and mandate) and criteria related to potential economic impacts (e.g., expected economic benefits).

Some of the criteria are also related to the "market" for the R&D—e.g., existence of a constituency that supports the work, potential for technology transfer.

Each project is rated high, medium, or low for each of the eight criteria. All of the criteria are weighted equally, and an overall score is derived for each project. It is intended that the lowest rated projects will no longer be supported. An important feature of this method is that the lab directors review the projects together and arrive at the ratings for each project by consensus.

3.7 Other Methods Which May be Useful in Certain Situations

Before discussing other methods, it should be noted that the above discussion of impact assessment methods simplifies the true situation considerably. First, as already noted, these methods are rarely used on their own—in any given assessment the most common practice is for two or more of these methods to be used in combination. Second, there are many other methods which are themselves combinations of the methods described above. For example, with regard to the use of the more quantitative assessment techniques, the dominant practice is to use data collected through interview, questionnaire, and case study exercises, and then evaluate that data within a benefit-cost methodological framework, an economic surplus framework, or an econometric exercise. (These latter two methods are discussed in Appendix A.) For example, the National Institute of Standards in Technology in the United States has carried out a number of studies which have involved assessing survey and case study data on the basis of benefit-cost principles in order to evaluate the economic and social impacts of its industry research support programs and its standards development activities.

Following are brief descriptions of five impact assessment methods which may be useful in certain situations. The first two are most applicable to R&D which has been carried out in the past, while the latter three are most useful for assessing R&D being considered for the future.

Patent Analysis

The principle behind patent analysis is that the technological performance of an R&D activity can be assessed by counting the number of patented products, processes, or systems which result from the activity. There are numerous problems with this technique, which are discussed in Appendix A. One of the main ones is that patent analysis provides no indication of whether the patented item is in use, who the users are, or how large the user group may be. However, with the new scheme to impose maintenance fees for patents, this method could become more predictive of use, and, therefore, more useful as an indicator of economic impacts.

Economic Surplus Methods

These methods have been used extensively in Europe to measure the "spin-off effects" of large R&D programs, such as the activities of CERN and of the European Space Agency. These R&D programs may generate spin-off effects by virtue of the fact that a major part of the scientific equipment necessary to carry out the R&D is purchased from industry. In many cases, the delivery of this equipment may represent a technical challenge to industry, the solution to

which may generate additional economic opportunities for industry by improving its technological know-how.

This technique is designed to determine the economic impacts resulting from this improved technological know-how, such as increased sales or increased employment. The data are collected through surveys and interviews of firms associated with R&D programs and institutes. The firms are asked to determine what part of their current sales are a result of the influence of doing work for the program/institute, and the resulting estimates are then applied to income and employment multipliers to derive an estimate of the total income and employment associated with the investment in R&D. While there are a number of problems with this method (discussed in Appendix A) there are certain aspects of the method that could prove useful if one is interested in estimating the economic impacts of these sorts of spin-off effects.

Mathematical Programming. The mathematical programming approach is similar to the integrated partial indicators approach in that weights are placed on a set of criteria. However, mathematical programming provides a more powerful and sophisticated priority setting technique in that it relies on the mathematical optimization of a multiple-goal objective function, subject to a resource constraint (available funding and human resources) to select a portfolio of research projects. As a consequence, this procedure has an advantage over the integrated partial indicators approach in that it selects an "optimal" portfolio, taking into account the various evaluation criteria and constraints imposed in the programming problem, rather than simply ranking research areas.

Mathematical programming provides a useful alternative for selecting research projects. However, like the integrated partial indicators approach, it is not particularly useful for evaluating too diverse a set of R&D projects. Also, if either the criteria for project assessment or the constraints faced in executing R&D projects (especially over time) are not well defined, there is a risk than an "optimal" but nonsensical solution can result.

The mathematical programming approach may not be especially useful for assessing the impacts and benefits of future R&D. Unless the connection between a set of projects and a certain magnitude of benefits is clear, it is not readily apparent how one would set up the objective function to be optimized. However, if the connection between a project and its potential for benefits can be specified, then this approach can be used to select a suite of projects that will maximize the potential for R&D investments to generate benefits. In sectors such as agriculture where the relationship between R&D and economic effects (such as changes in productivity) is clearer and more direct, then it may be possible to incorporate a specific economic impact in the objective function (such as a certain percentage change in input productivity or yield rates) and then use the mathematical programming approach to select the "optimal" suite of R&D projects. (It should be noted that, in an example such as this, the distinction between mathematical programming and simulation models discussed below becomes somewhat blurred.)

Cost-Effectiveness Analysis. Cost-effectiveness analysis is a variant of benefit-cost analysis which is applicable in cases where it is not feasible to assess the potential benefits of a project. A common example of the types of projects that fall into this category include R&D related to

standards development and policy objectives, such as improved health and safety. For example, the cost-effectiveness approach is the recommended method for determining resource allocations within the Australian National Standards Laboratory.[9] It is a comparatively simple task to postulate that R&D in these areas will be of benefit to society. However, in general, it is not a simple task to quantify these benefits, especially for future R&D. However, if the benefits are well defined and accepted as the objective of the R&D, then the R&D project choices that have to be made involve decisions regarding alternative ways of reaching the objective. In essence, in cost-effectiveness analysis one employs benefit-cost principles to evaluate the costs of alternative projects which, at least conceptually, are capable of meeting the stated objective. The decision rule then becomes one of selecting the project that can attain the objective at the least cost to society. The difficulty with this approach is that one must ensure that the alternative projects under consideration have the potential to generate the same kinds of societal benefits (i.e., that they are equally capable of reaching the stated objective and that one choice will not yield significantly greater or smaller benefits than the other choices under consideration).

Simulation/Econometric Models. The use of simulation/econometric models to gauge the impacts of R&D can be traced back to the 1950s and the pioneering work of Abramowitz (1956), Solow (1956), and Griliches (1958).[10] Simulation models provide a framework within which the impact of R&D projects on a specified objective function(s) can be gauged. These models can vary in complexity from simple spreadsheet-based functions to elaborate econometric models. While the simulation/econometric models that have been used to assess R&D differ in their construction, a number have been based on a production function approach where the impact of R&D can be gauged by changes in factor productivity. There are some significant technical problems with this approach (especially with the more formal econometric models), which are discussed in Appendix A.

Simulation models which overcome some of these difficulties and are more applicable to priority setting and project evaluation have recently been developed, although predominantly within the agricultural R&D sector where the connections between R&D and sectoral economic impacts are clear and more direct.[11] Within these models various agricultural demand, supply, price and production relationships, financial constraints, and technical relationships (incorporating

---

[9]   *Op. cit.*, BIE, 1992, pg. 43.

[10]   Abramowitz, M. (1956), *Resource and Output Trends in the United States, American Economic Review Papers and Proceedings*; Solow, R. (1957), *Technical Change and the Aggregate Production Function, Review of Economics and Statistics*; Griliches, Z. (1958), *Research Costs and Social Returns: Hybrid Corn and Related Innovations, Journal of Political Economy*.

[11]   An interesting example, noted in the literature, of an agricultural sector simulation model is the Pinstrup-Anderson/Franklin model (Pinstrup-Anderson, P. and Franklin, D (1977), *A Systems Approach to Agricultural Research Resource Allocation in Developing Countries, Resource Allocation and Productivity in National and International Research*, ed. T.M. Arndt, D.G. Dalrymple, and V.W. Ruttan, University of Minnesota Press, Minneapolis).

assumptions regarding the probability of research success and adoption) are specified. The sectoral simulation models are then used to evaluate and rank the costs of alternative R&D projects and their contributions to certain sector-specific goals (such as increased crop productivity).

The advantage of simulation models is their flexibility. They can be constructed as relatively simple or as extremely complex models; they can incorporate ranking or optimization algorithms; and they can readily include probabilistic functions. Their major disadvantage is that, to be at all useful, they must accurately reflect the relationships between technological advancement and economic development. This generally requires the construction of a relatively complex model which typically involves extensive amounts of time to construct and data to operate. As noted above, in sectors such as agriculture where the relationship between R&D and the resulting economic consequences is fairly clear and direct, the development of a roughly accurate model may not be as problematic. However, as noted in Appendix A, for more "complex" sectors and less applied types of R&D the relationships between economic benefits and R&D are often indirect, non-linear, and exhibit markedly different lag distribution patterns. In these instances the development of simulation/econometric models becomes quite burdensome in terms of the development effort and data required, and the utility of the resulting model is generally quite suspect.

3.8   Summary

Exhibit 3.1 summarizes the applicability and the most important characteristics of the impact assessment methods discussed in Sections 3.1-3.6.

Exhibits 3.2 and 3.3 present essentially the same information (except for the information on strengths and weaknesses and relative costs) from a different perspective—for each time frame, type, and purpose of R&D they list the impact assessment methods which are potentially useful.

The overall conclusion to be drawn from these sections is that there are methods available which can provide at least a partial picture of the impacts of government R&D activities and, in many cases, can provide reasonably accurate quantitative estimates. The main gaps that remain are methods for assessing the impacts of basic/strategic R&D (for obvious reasons) and methods for providing quantitative estimates of the impacts of R&D intended to provide information needed for the development of government policies and regulations—although case studies can be used to illustrate these impacts and even benefit-cost analysis can be used in certain circumstances.

At first glance the methods discussed in this section do not appear to be much different from methods that were used 15-20 years ago—at least the names are pretty much the same. There have been many important advances, however, such as the following:

- modified peer review—the element of expert socioeconomic opinion has been added to expert scientific opinion, so that the peer review approach can now provide information on both scientific quality and relevance/usefulness;

Exhibit 3.1: Applicability and Characteristics of Different Methods

Methods	R&D Time Frame	R&D Type	R&D Purpose	Strengths and Weaknesses	Relative Cost
Modified Peer Review	Past, on-going, and future	All	All - although least well suited for Category 1	**Strengths** • Relatively easy to organize. • Can provide valuable information on potential impacts • Probably the best method for basic/strategic R&D. **Weaknesses** • Relies on the opinions of a small number of people. • Qualitative information only.	Low/Medium
User Surveys	Past and on-going	Applied research and development	Category 2 to Category 4	**Strengths** • Overcomes the problem of a small number of respondents. • Possible to develop quantitative indices. **Weaknesses** • Structuring the survey and analyzing the results can be tricky.	Medium (often requires considerable time to identify users, develop survey methodology, and analyze results)
Benefit-Cost Methods	Past (can be used for on-going and future R&D in certain circumstances)	Applied research and development	Category 4 (can be used for Category 2 and Category 3 R&D in certain circumstances)	**Strengths** • Can provide reasonable and defensible estimates of potential benefits. • Provides a structured framework for assessing R&D projects which forces the right questions to be asked. **Weaknesses** • Can be very time consuming and labour intensive. • Results are critically dependent on assumptions which can be highly uncertain. • Because of cost and time requirements can only be used for a limited number of projects.	High (data collection requirements are very demanding)

R&D types: Basic/strategic research, applied research, product/process development
R&D purposes: Category 1 - R&D infrastructure, Category 2 - Policy development, Category 3 - Policy attainment, Category 4 - Industrial innovation

Exhibit 3.1: Applicability and Characteristics of Different Methods

Methods	R&D Time Frame	R&D Type	R&D Purpose	Strengths and Weaknesses	Relative Cost
Case studies	Past	Applied research and development	Category 2 to Category 4	**Strengths** • Can provide good illustrations of the relationship between R&D and its impacts. • Probably the best method, in general, for Category 2 R&D. **Weaknesses** • Generally there is no way to "add up" the results of a group of case studies to obtain a measure of the total impacts of the group. • The results cannot be extrapolated to other R&D projects which are not in the group.	Medium (depending on the number of case studies)
Partial Indicators	Past and on-going (and future to a limited extent)	All	All	**Strengths** • The information required to specify the indicators is relatively easy to collect. • Probably the best method for on-going monitoring. **Weaknesses** • The individual indicators can generally only be "added up" on a judgemental basis, making overall impact assessment more difficult. • Provides only a very partial picture of impacts.	Low
Integrated Partial Indicators	Future	Applied research and development	Category 2 to Category 4	**Strengths** • An easy but structured of way to identify research priorities. • Forces the decision makers to explicitly consider the key determinants of impacts. **Weaknesses** • Totally relies on the judgement of (usually a few) individuals. • Potential for bias in assigning weights to different criteria.	Low

R&D types: Basic/strategic research, applied research, product/process development
R&D purposes: Category 1 - R&D infrastructure, Category 2 - Policy development, Category 3 - Policy attainment, Category 4 - Industrial innovation

Exhibit 3.2:   Methods Useful for Assessments of Past R&D

| R&D Type | R&D Purpose | | | |
	Category 1 R&D Infrastructure	Category 2 Policy Development	Category 3 Policy Attainment	Category 4 Industrial Development
Basic/Strategic	(Modified Peer) (Partial Indicators)	Modified Peer (Partial Indicators)	Modified Peer (Partial Indicators)	Modified Peer (Partial Indicators)
Applied	(Modified Peer) (Case Studies) (Partial Indicators)	Modified Peer User Surveys Case Studies (Benefit-Cost) (Partial Indicators)	Modified Peer User Surveys Case Studies (Benefit-Cost) (Partial Indicators)	Modified Peer User Surveys Benefit-Cost Case Studies (Partial Indicators)
Development	(Modified Peer), (Case Studies) (Partial Indicators)	Modified Peer User Surveys Case Studies (Benefit-Cost) (Partial Indicators)	Modified Peer User Surveys Case Studies (Benefit-Cost) (Partial Indicators)	Modified Peer User Surveys Benefit-Cost Case Studies (Partial Indicators)

( )   =   Method is of some value in certain circumstances

R&D Type	R&D Purpose			
	Category 1 R&D Infrastructure	Category 2 Policy Development	Category 3 Policy Attainment	Category 4 Industrial Development
Basic/Strategic	Partial Indicators (Modified Peer)	Modified Peer Partial Indicators	Modified Peer Partial Indicators	Modified Peer Partial Indicators
Applied	Partial Indicators (Modified Peer)	Modified Peer Partial Indicators Integrated Indicators (User Surveys)	Modified Peer User Surveys Partial Indicators Integrated Indicators	Modified Peer User Surveys Partial Indicators Integrated Indicators (Benefit-Cost)
Development	Partial Indicators (Modified Peer)	Modified Peer Partial Indicators Integrated Indicators (User Surveys)	Modified Peer User Surveys Partial Indicators Integrated Indicators	Modified Peer User Surveys Partial Indicators Integrated Indicators (Benefit-Cost)

( ) = Method is of some value in certain circumstances

- user surveys—methods have now been developed for asking better questions, better interpreting the results, and ensuring a higher degree of validity;

- benefit-cost analysis—this is one technique that was thought to be totally impractical for assessing R&D 10-15 years ago; the technique has been considerably refined since the mid-1980s, and some important new approaches have been developed, such as the concept of partial benefit-cost analysis;

- partial indicators and integrated partial indicators—the process for applying these methods has been improved (e.g., the importance of heavily involving the research managers is now widely recognized); also, these techniques have begun to include consideration of the "receptor capacity" and the market for the R&D as important indicators.

The following discussion of the information which is taken into consideration in each of the methods to draw conclusions about impacts is intended to provide some additional insights regarding these methods.
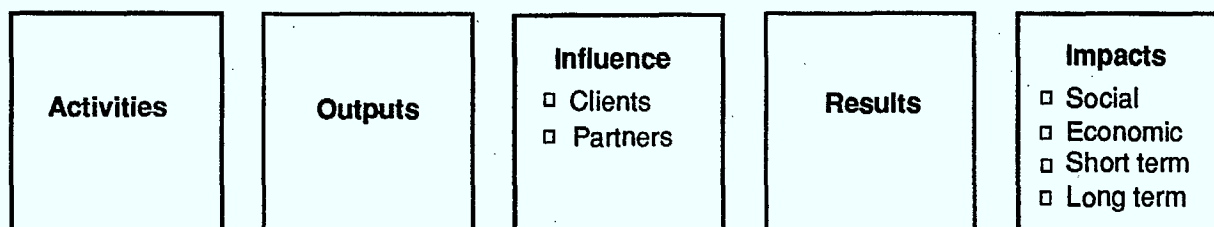
Exhibit 3.4 shows at the top the central portion of NRC's "model for performance framework" and at the bottom the main impact assessment methods together with the time frame of the R&D to which they are normally applied. The part of the performance framework model covered by the bracket shows the information that is taken into consideration by each method.

For example, the methods which are applied to on-going and future R&D—modified peer review, user surveys, and integrated partial indicators—have the most distant perspectives of all the methods, since the only "hard" information available is information on activities (and preliminary outputs for on-going R&D). Information on influence is brought into these assessments via the knowledge of participants in the process—the socioeconomic experts for modified peer reviews, the potential users for user surveys, and external advisers (usually) for integrated partial indicators. (This is the missing element in traditional peer review, and the reason traditional peer review cannot be used for socioeconomic impact assessment—see Appendix A.)

When these methods are applied to past R&D, the assessments of impacts are at least somewhat less speculative, since the results of the R&D are in-hand. In user surveys, for example, the respondents should be able to comment fairly knowledgeably about the likely usefulness of the results for their organization. Downstream impacts can then be inferred based on the logical linkages between user effects and these impacts.

Benefit-cost methods focus on the impact end of the performance model. Costs need to be determined for all stages up to benefits, and the attribution of benefits to the R&D can only be made by tracking from activity through intermediate steps to impact. Benefits are measured directly, not from a distance, and should be more quantified than other methods, even if not directly and clearly attributable or accurate. For these reasons this approach is often more credible among readers (if not always so among practitioners or experts).

Exhibit 3.4:   Performance Framework Model

Activities	Outputs	Influence □ Clients □ Partners	Results	Impacts □ Social □ Economic □ Short term □ Long term

Modified peer review and
user surveys - ongoing and
future R&D

Modified peer review and user
surveys - past R&D

Benefit cost methods and case
studies - past R&D

Integrated partial indicators - future R&D

Partial indicators - past, on-going, and future R&D

Case studies have essentially the same characteristics—they describe a chain of events linking activities through to impacts. The difference is that the emphasis is generally less economic, with benefits being described in qualitative and behaviour influencing terms.

Partial indicators can operate at any—or all— parts of the model. Impact indicators tend to focus on key attributes which are found to be essential for maximizing impact. They can be related to activities or outputs if these specific activities or outputs have been shown to be or planned to be key to benefits. Other indicators can be closer to the impact end of the model, and can relate to client influence, or specific targeted short or longer term impacts. They can be broadly based or not, depending on the approach taken and agreed to by managers setting performance objectives.

4.0 Methodologies for Other S&T Activities

4.1   Background

Research and development activities are generally considered to be at the center of a broader group of activities known as science and technology (S&T). There are a number of different interpretations of the meaning of these terms, and to minimize confusion and maximize compatibility with other literature, the definitions of the Frascati Manual as incorporated by the OECD and Statistics Canada have been accepted for this report. Research and development is defined as "creative work undertaken on a systematic basis to increase scientific and technical knowledge and to use this knowledge in new applications". There is also within the S&T portfolio a set of "Related Scientific Activities" (RSA) which includes a number of categories which are in direct support of R&D. The specific categories of activities included within R&D and RSA are given in Exhibit 4.1.

Most S&T activities can be reasonably placed within this structure, however certain cases require some interpretation. For example, funding of graduate or post graduate activities can fall within R&D grants and contributions, research fellowships or education depending on whether the primary purpose of the expenditure is to produce R&D results or to educate and train individuals.

4.2   Impact Assessment of Other S&T Activities

The literature search undertaken for this study revealed that while there is considerable interest and research on impacts of R&D activities, there is relatively little effort on other S&T activities, with the possible exception of training and education . Since much RSA is in support of R&D, the linkage with ultimate economic and social impacts is even more diffused than for R&D, and consequently may be more difficult to track. The information provided by departments both through reports and discussions in the workshops also tended to focus on R&D activities, reflecting departmental priorities and capabilities.

In fact, some types of R&D as defined in Exhibit 4.1 are also outside the main discussion areas of this report, and could be included within the category of other S&T activities. These include supporting contracts and research fellowships.

As is the case with some R&D activities, intermediate impacts such as influence and affect on the recipient can be often be more directly attributed to other S&T activities, and impact assessment may focus at this level. Intermediate analysis of this nature can be supplemented by logically inferring and attributing related longer term impacts.

Exhibit 4.1: Categories of S&T Activities

Research and Development	
Category	**Description**

Departmental
    Intramural
    Collaborative
    Contract-in
	Performed within department or program

Contracts
    R&D
    Support
	Contracted R&D, Support to Departmental R&D

Grants and Contributions	Provided to benefit recipient
Research Fellowships	Fund advanced research training

Related Scientific Activities	
Category	**Description**
Scientific Data Collection	Geological surveys, etc...
Information Services	Recording and disseminating STI libraries, journals, S&T Advisory Services
Testing and Standards Development	Calibrations development of new standards
Feasibility Studies	Engineering studies, demonstration projects
Museum Services	Collection, display of natural phenomena
Education Support Grants	Post secondary S&T education

Note: Additional S&T categories for both R&D and RSA include Administration of Extramural Funding and Capital Expenditures

Although the literature is less advanced, it is clear from the close relationship between the impact pathways for most R&D and other S&T activities, that many of the methodologies appropriate for assessing impacts of R&D are also appropriate for other S&T activities. These include modified peer review, client and user surveys, benefit cost, case studies and partial indicators of impacts. These approaches are summarized in the preceding chapter, and presented in more detail in Working Paper #1 of this study.

In certain cases, alternative approaches may need to be considered, and there may be a need to modify or adapt specific approaches to the differences in linkages and attribution between activities and possible impacts. For example, some of the activities in both the R&D and RSA categories, such as information services and administration of extramural funding (contracts and grants and contributions) have a more direct service aspect than other activities which directly perform R&D. In these cases, there can be even less control over ultimate impacts than for other activity classifications. For these cases, it may be appropriate to make use of a user survey and consider the relationship between impact assessment and service characteristics.

As is the case for R&D, each specific methodology has a varying degree of applicability to the categories of RSA, depending on the specific nature of the activity or program. Exhibit 4.2 presents a summary linking RSA activities to methodologies. The large majority of RSA can be assessed using these approaches.

Exhibit 4.2:   Methods Useful for Assessing Other S&T Activities

---

Related Scientific Activities	
Category	**Methodologies**
Scientific Data Collection	All*
Information Services	All*
Testing and Standards Development	All*
Feasibility Studies	All, but Peer Review
Education Grants	All*

*All: Modified Peer Review, User Surveys, Benefit-Cost, Case Studies, Partial Indicators, Integrated Partial Indicators

---

5.0 Other Findings

5.1   Increasing Emphasis On Impact Assessment

It is clear from the review of practices in other countries and in Canada that there is an increasing emphasis on assessing the impacts of government R&D, both in other countries and in Canada. Several countries were identified in which impact assessment systems have been set-up relatively recently. For example, in Sweden the Swedish National Board for Technical Development is in the process of developing and implementing on-going performance monitoring systems, which include the assessment of impacts or potential impacts for all of its R&D programs (approximately 100 programs). They were directed to this by the Swedish Parliament, but it is also the case that the performance monitoring systems have been strongly supported by the R&D managers, because they perceive that these systems will assist them with program management. Similarly in Australia, the act of Parliament governing the CSIRO was recently amended to specifically direct that retrospective, on-going, and prospective assessment of programs and activities become part of the management process. As a consequence, CSIRO has set up a fairly elaborate system for picking priority research fields, partly on the basis of factors related to R&D (the payoff and Australia's capability), but equally importantly on the basis of the potential impacts of the R&D (both the benefits from the R&D and Australia's ability to capture these benefits). The review identified no example of a country which is doing less impact assessment currently than it was a few years ago.

The same situation prevails within the Canadian government. As is clear from Section 3.0, a number of departments are doing more impact assessment that they were previously. Some departments which up until recently have not been especially active in this area have become active, and some departments that have been active in the past have become even more active.

This increasing emphasis within Canadian government departments is not just occurring as a result of the influence of central agencies or program evaluation groups. In fact, it appears to be driven primarily by the research managers themselves. This is certainly the case in Agriculture Canada, where the emphasis on assessing the economic impacts of future R&D has bean due, to a large extent, to the influence of the senior management of the Research Branch. Other examples of the same phenomenon are described in Section 3.0. For example, the Portfolio Analysis system recently implemented by NRC's Institute for Environmental Chemistry was developed and implemented by the managers of the Institute. Similarly, the rationalization exercise currently being carried out at the DOC research laboratories has been spearheaded by the senior lab managers.

The main reason for this increasing emphasis is the budgetary pressure currently being experienced by most governments. Governments are asking many more questions than they did in the past about what they're getting from their expenditures on R&D. Therefore, there is a need for R&D program managers to be able to document the impacts of past R&D in order to explain and defend their programs, as well as manage them. In addition, program managers

recognize that the impacts which result from the R&D will be used as one of the main measures by which their programs are judged in the future. Therefore, there is a greatly increased interest in the use of prospective impact assessment methods for selecting R&D projects and project areas and for monitoring on-going projects.

5.2 Areas of Increased Emphasis Within Impact Assessment

In addition to increasing emphasis on impact assessment in general, it appears that there are two significant changes occurring within the impact assessment practices of a number of countries:

- there is increasing emphasis on prospective impact assessment, especially for future R&D;

- there is increasing emphasis on more careful consideration of the "receptor capacity" of the organizations who are intended to be the primary users of the R&D.

The increased emphasis on prospective impact assessment is seen in both in other countries (e.g., Sweden and Australia) and in Canadian departments and agencies. Some examples in Canada include:

- Agriculture - the use of a return on investment model as part of the project selection process;

- DOC - the planned use of criteria related to potential economic and social impacts for evaluating new projects;

- CANMET - the inclusion of the assessment of future economic benefits in the performance indicator system for on-going PERD projects oriented toward economic development.

The reason for this trend is discussed in Section 5.1.

The term "receptor capacity" means the level of interest in the R&D by the organizations which are its primary intended users and the capability of these organizations to actually apply the R&D. The consideration of receptor capacity is a part of many of the prospective assessment methodologies used in other countries which we reviewed—see, e.g., the description of the CSIRO system in Section 3.6.

Receptor capacity is also being increasingly considered as an important factor in R&D planning within the Canadian government. For example:

- In Agriculture Canada, the prospective assessment methods being implemented include an analysis of the market for the products and processes that are expected to result from the R&D, as well an analysis of the extent to which industry is likely to

take up the R&D. The formal analysis of the potential market for the use of R&D results is a new element in the Research Branch's planning process.

- The criteria used in DOC's rationalization exercise include some related to the "market" for the R&D—e.g., existence of a constituency that supports the work, potential for technology transfer. In addition, DOC is now obtaining opinions from their intended clients regarding the relevance of their on-going work. This includes intended clients within the Department for their policy-oriented research.

- Within Energy, Mines and Resources, the performance indicator system for on-going PERD projects includes the analysis of the extent to which realistic uses and users for the research have been identified. Also, it was clear from the discussions at the CANMET workshop that CANMET will be giving much greater consideration to all the factors required for commercialization of the R&D, including factors that are beyond the control of the Department.

- The review of potential projects within NSERC's Cooperative Research and Development program includes an analysis of the potential for Canadian companies to capture the benefits of the research. The receptor capacity of Canadian industry was also one of the criteria used (at least theoretically) in the selection of the networks in the Networks of Centres of Excellence program.

- NRC now includes the "influence" of its R&D programs as one of the key elements of the program logic. The recent strategic planning studies for the Engineering Sector and the Biotechnology Program included assessments of the receptor capacity for the R&D and recommendations regarding which industrial sectors the R&D should focus on in order to achieve the greatest impact.

The inclusion of receptor capacity in prospective impact assessment methods is interesting, because, for economic development oriented R&D, it implies the need to understand the country's industrial structure. As discussed below, the people responsible for carrying out the assessments may not always have this understanding. This was one of the problems, for example, with the application of the receptor capacity criterion in the selection process for the Networks of Centres of Excellence Program.[1]

5.3   Sufficiency of Information Currently Available to Canadian Departments/Agencies

Four of the six departments which participated in this study indicated the need for additional information on the impacts or potential impacts of its R&D activities. For one of the other two, NSERC, this subject was not discussed. It should be noted, however, that NSERC has been very active in conducting impact analyses for a number of years, and they are currently planning

---

[1]   *Interim Evaluation of the Networks of Centres of Excellence Program*, The ARA Consulting Group, February, 1993, Section 8.0.

to implement an on-going performance measurement system which will include additional information on the impacts of their programs. In the other department, DOE, there was no clear consensus regarding whether the available information on impacts and potential impacts of their R&D activities is sufficient for their purposes. Some people thought that the available information is sufficient, while others thought it would be useful to have more information on the uses of past research and the intended uses of on-going and planned research.

Our conclusions regarding the opinions of the other four departments, as inferred from the group meetings, are summarized below:

- Agriculture - It is felt that more complete information on the economic impacts of past projects would be useful for planning purposes. Also, at least for some parts of the organization, more information on the perceived importance and usefulness of the research by industry would be useful for better responding to client needs and defending programs. In addition, the establishment of a formal performance monitoring system which includes some indicators related to impacts and the work that is being done on the development of the return on investment methodology indicates that there is a perceived need for more formalized and standardized methods for both on-going performance monitoring and project selection.

- DOC - There seemed to be a general consensus that the kind of system planned for the PERD program for assessments of past research would provide useful information for defending programs, and the kind of system used by CSIRO for selecting project areas might also be useful. (The latter is probably not too dissimilar from the selection procedure the labs are planning to implement which will use the criteria currently being used in the rationalization exercise.)

- CANMET/GSC - Within CANMET, there is a perceived need for more information both on the economic impacts of the R&D (specifically, the impacts on wealth creation) and the social impacts of the R&D (specifically, the extent to which R&D results provide support for government policies and priorities). In addition, there appeared to be a consensus that more information is required for making project decisions—in particular, information on potential impacts measured in quantitative terms for large projects and indicators of potential impacts (such as measures of client need and measures of client receptor capability) for smaller projects. Within the GSC it was agreed at the February 3 performance measurement seminar that there is a need for better measures of performance (and most of the possible measures discussed were measures of impacts).

- NRC - It appears that there is sufficiently good information available on the impacts of past projects and programs for use in defending programs. There does appear to be a shortage of information, however, on the potential impacts of on-going and possible future projects/project areas. Both the Engineering Sector strategic assessment and the strategic assessment of the Biotechnology Program call for choosing R&D projects so as to make the maximum possible contribution to Canadian

industrial competitiveness. This implies the need for assessing the potential impacts of future R&D (including assessing the market for the R&D results). Although this is done to some extent informally at the present time, the attendees at the meeting felt there was a need to formalize and standardize the methods used. It was also noted that economic analyses of potential impacts of on-going and future projects are being requested by the various NRC advisory committees on an increasingly frequent basis.

Generally it appears that there is a need for more information, especially regarding the potential impacts of on-going and future projects.

5.4   Potential Challenges

Interest in Impact Assessment. Based on the activities carried out in this study, it was possible to reach conclusions regarding the level of interest of the participating departments in impact assessment for five of the six departments. Four of these have a high level of interest, and one has a medium level of interest. Specific comments relating to each department are:

- Agriculture - There is a very strong interest by the senior management of the Research Branch in setting up systems for formally assessing the economic and social impacts of R&D projects—both work carried out in the past, on-going projects, and, especially, possible future projects.

- DOC - The fact that there is at least a medium level of interest in impact assessment can be inferred from the implementation of the rationalization exercise. Also, past impact assessment work has not been viewed negatively by the labs.

- CANMET/GSC - There is a high level of interest among the senior managers in CANMET and in the GSC in assessments of impacts and potential impacts of their R&D projects. A number of key points were agreed upon in the CANMET workshop: for example, the need for formalized criteria involving impact considerations for project selection and the need for post-mortem assessments of all completed projects against their socio-economic objectives (see Appendix B for details). In the GSC Performance Measurement Seminar strong interest was expressed in having more and better measurements of impacts for priority setting and resource allocation, on-going monitoring, and defending programs.

- DOE - There is some degree of interest within this department in assessing the socioeconomic impacts of its R&D activities; however, the full extent could not be determined within the limited scope of this study. The department is active in socioeconomic impact assessment in other (non-R&D) areas—for example, the estimation of the economic spin-offs of its parks (which it considers as living laboratories), the assessment of socioeconomic impacts of regulations by Conservation and Protection, the study of the economic consequences of global warming by AES. These methods have not yet been applied to its R&D activities specifically on a large scale.

- NSERC - The level of interest in impact assessment is fairly high throughout the organization—i.e., not just in the evaluation and policy and planning groups, but also among the program managers.

- NRC - There is clearly a very high level of interest within NRC in impact assessment, from the President on down.

In general, the level of interest within Canadian government departments and agencies in assessing the impacts or potential impacts of their R&D activities does not seem to pose a problem.

Capability. The capability to carry out these impact assessments, on the other hand, may be a slight problem. All six of the departments indicated that they felt they needed additional assistance and/or resources to carry out the kinds of impact assessments that would be desirable. Some of these departments have already arranged for the additional expertise required (or have had such arrangements in place for some time), but for others additional arrangements would have to be made. For example, at the CANMET workshop there was a recognition that in order to set up the impact assessment systems the attendees felt are desirable, greater expertise is needed than is currently available—both increased internal expertise and external expertise. With regard to internal expertise, it was suggested that the research managers need training in assessment techniques in order to fully understand the implications of these techniques and incorporate them in their management and monitoring procedures. It was also agreed that formal assessments of socioeconomic impacts should be done by specialist groups, working together with the research managers. Similarly, at the NRC meeting, there were some opinions expressed that additional expertise and resources would be required to carry out some of the kinds of impact assessments that were discussed. In particular, the individual institute officials made it clear that they don't have the expertise to do the industrial/economic analysis that is required for credible assessments of either past or potential future economic impacts. On the other hand, they indicated that they didn't want impact assessment done "on them" by outsiders.

This last point made at the NRC meeting is important. The NRC institute officials reached the same conclusion that is apparent from the literature—impact assessment methods are more effective and better accepted if the senior research managers are full participants in the process. The CSIRO system for assessing future research is a good example, as is the DOC rationalization exercise.

Limited internal capability in this area is not a major problem. It is true that the research managers need some external advice and assistance in order to carry out impact assessments, but probably much less assistance than they think. It is almost certain that if they and the researchers can put in the time in collecting the data and documenting the impacts, then the amount of external assistance required is relatively small. CSIRO is a good example—the assessments are essentially made by the lab directors, with some but apparently not much input from expert advisors regarding such things as industrial structure and level of industrial interest and capability. The GSC Performance Measurement Seminar provides another good example. In the morning session the attendees identified the lack of sufficient internal expertise as one of

the potential barriers to getting better performance measures. On the other hand, in the afternoon session the attendees were given five typical GSC projects as case studies and, entirely on their own, they were able to develop a reasonable set of performance measures for each of these projects.

Data Collection and Documentation. One very important finding emerged from the literature review which applies to virtually all of the methodologies discussed in this report. Namely, many of the failings of current methods result not so much from internal methodological problems as from attempts to apply these methods in a climate characterized by poor project monitoring, evaluation planning, and data collection practices. That is, it is often the case that the objectives of the R&D are not clear and the data to evaluate the socioeconomic impacts are not available and cannot be obtained.

The literature indicates that in order for effective assessment of economic and social impacts to be possible, R&D planners, in conjunction with independent experts where necessary, have to more specifically identify how a particular proposal will serve the mandate of the R&D funding organization. In addition, individual researchers need to get actively involved in explicitly defining the purpose of the R&D and the range of anticipated results, and they must define clear testable objectives against which assessments of the progress and eventual success of the R&D can be made. The purpose of the R&D and its objectives should then be linked directly to the evaluation process, and appropriate monitoring and data collection activities should be established at the outset to provide the information necessary to assess the extent to which the R&D activity has achieved its objectives.

This same situation clearly prevails in Canada and is one of the main obstacles to the implementation of good impact assessment practice for R&D activities. In our meetings with the individual departments and agencies, it became clear that more impact assessment work is being carried out than is readily apparent. It's just that the work is being carried out fairly informally in many cases, the results of the assessments are not documented, and very little data are maintained upon which future assessments can be based. For example, in the NRC workshop, participants at first said they didn't do much in the way of impact assessments for possible future projects and project areas. However, when the CSIRO system was discussed, attendees agreed that their institutes do carry out this kind of assessment—including the assessment of the "attractiveness factors"—as part of their normal business planning. The problem is that the assessments are sometimes not very rigorous, and they are rarely documented explicitly. Another example arose in the workshop with DOE officials. As noted above, there was some interest expressed in documenting the uses of R&D that has been carried out in the past and the intended uses for planned and on-going R&D. It appeared that this would not be difficult to do (at least for some of the R&D), but it is not done currently.

Lack of uniform knowledge regarding impact assessment methods. Even in the limited review of the practices of Canadian government departments and agencies carried out as part of this study, it became clear that there is a considerable lack of uniformity in the methods and terminology used. For example:

- Some user surveys base their samples on clients, while others base their samples on beneficiaries (see Section 3.2 for definitions of these terms). Also, many of the questions are asked in very different ways, and there are differences in the amount and types of information supplied to the reviewers.

- There are some significant differences in the ways in which benefit-cost methods are used—e.g., differences in deciding which costs should be included, different ways of handling the issue of incrementality and attribution.

- There is no uniformity at all with regard to the type of information that is presented in case studies.

These are just some obvious examples. There are many more.

Non-uniformity is not necessarily a bad thing. It would be incorrect to suggest that all government R&D organizations should use the same standard set of impact assessment methods in exactly the same way. Different organizations have different information needs. For example, one organization might be more interested in the views of its direct clients than the views of its beneficiaries, while the opposite might be the case for another organization. Also, there are always specific questions of interest to individual organizations that are irrelevant or meaningless to others.

However, in cases in which the non-uniformity is caused by lack of information about what methods are available, their strengths and weaknesses, and how they should be applied to provide information that is valid and as useful as possible, it can be a bad thing. The following problems can occur as a result:

- mistakes can be made in the application of a method which can invalidate the results;

- a sub-optimal method may be used—or an appropriate method may be used sub-optimally—resulting in less useful information than would be the case otherwise;

- two different studies intended to produce comparable results may not do so;

- inconsistencies in the use of terminology may lead to serious confusion.

The challenge to be addressed is not uniformity in application of impact assessment methodologies, but rather broadening the knowledge base of those performing impact assessment and sharing of that knowledge with those making use of this information for strategic, policy and operational management purposes.

In principle this should not be a difficult problem to solve. One could start fairly modestly by suggesting common terminology and providing additional information regarding how to implement the most common assessment methods (i.e., more information than provided in this

report). This would not be done in a prescriptive way—i.e., departments and agencies would still need to retain flexibility to adjust methods to suit their specific needs.

Conclusion. In spite of the challenges noted above, it can be concluded from this study that it is both feasible and useful to incorporate socioeconomic impact assessment as an element of management practice for government R&D activities. It needs to be kept in mind, however, that the appropriate amount and intensity of impact analysis, as well as the methods used, will vary with the type and purpose of the R&D and the information needs of the department or agency.