100448

International Development Research Centre

Centre de Recherches pour le developpement international

P.O. BOX/C.P. 8500
OTTAWA CANADA
K1G 3H9

TOWARDS GREATER EFFECTIVENESS OF INDUSTRIAL RESEARCH INSTITUTIONS: SOME TOOLS AND TRENDS

AMITAV RATH

April, 1994

CONTENTS

BACKGROUND

The International Development Research Centre, Canada, supported a research project that included seven case studies of selected Latin American industrial technological research institutions (ITRIs). The major objective of the ITRI project was to determine "how to make technological research institutions more *effective* in terms of providing relevant services and technological inputs to industry" (project document). The project also aimed to determine strategies for ITRIs to make the transition from their present conditions to more "effective" conditions. The project recognized that there was a new environment emerging, a "deregulated and market oriented environment," and wished to ensure the "sustainability of the ITRIs...in the context of the new environment, characterized by...reduced government expenditure" (project document).

From these objectives, a key operational issue for the entire project becomes the determination of the "effectiveness" of R&D. Not surprisingly, the question of the "effectiveness" of R&D has been with us for a long time, at least since the inception of organized R&D, and "has been the subject of the Industrial Research Institute programs since its foundation in 1938" (Hackett 1962). An early review of the issue of effectiveness is provided by Argyris (1968). We shall see that the question has become increasingly important for science policy researchers, government policy makers, and the managers and scientists working in such institutions since the 1980s. As resource demand for R&D has increased, the question has also assumed greater importance for the public, which ultimately provides the resources. Further, as newer theories in economics and business highlight the critical role of innovation for economic growth and competitiveness, the question of "effectiveness" has taken on additional significance.

Certainly, to devise any plan to reform an institution or system for greater effectiveness, we must first be able to clearly define what we mean by "effectiveness." What are its characteristics, and how can it be measured? Are the measures of effectiveness the same for different types of research institutions and their programs? Are some types of research and institutions more productive? If we can answer the above, then another set of questions looms. What inputs, organizational structures, processes, and management methods affect effectiveness? An ability to determine how different contextual and managerial factors affect the effectiveness of an institution can help policy makers, research managers, and researchers themselves to modify or remove the factors that reduce effectiveness and add or supplement those that increase it.

Inputs → | Type of ITRI and its Internal Dynamics | → Outputs

↑

Environmental Context

Figure 1

To define "effectiveness," we first set out a very simple schematic framework of an ITRI. Figure 1 serves to illustrate some of the key groups of variables that we may have to deal with to determine effectiveness and the factors that are likely to affect it. Simply put, effectiveness must depend on the characteristics of the output; the output, in turn, is likely to depend on the nature of inputs, the internal structure and dynamics of the institution, and the institution's environmental context.

In the next section, we discuss the nature and various roles of an ITRI, its structures, its inputs and outputs, the environmental context, and the interactions with users. In the second section, we discuss the question of effectiveness and suggest that it consists of at least three components — the efficiency of outputs, the appropriateness of the outputs, and, finally, the use and application of the outputs. In the third section, we discuss issues related to the evaluation. In the fourth section, we discuss what is known about the interrelationships between the different variables. In the final section, we review certain trends in public policy in a number of countries with respect to a restructuring of the ITRI and the innovation system more broadly.

1. GOALS, OBJECTIVES, AND CONTEXT

We start with the issue of the effectiveness of research institutions. To determine if any institution is effective, we first must know its objectives or mission. Every laboratory is created with a particular mission in mind. Public research laboratories are of many different types. They can be distinguished by their functions: laboratories involved in regulation, control, and standards setting for the government; project management laboratories that supervise the design and building of technological systems such as telecommunications, transport, and others; service laboratories that undertake mainly contract work; science frontier laboratories that are engaged in the development of new scientific knowledge and training of scientists; production-oriented facilities that may specialize in building prototypes and test equipment; and so on (see Miller 1992, p. 9; Clarke and Reavley 1988, p. 34; and OECD 1987a for some typologies and discussions of different types of research institutions). Further, the research institutions may also be sectorally defined in terms of health, agriculture, industry, aerospace, energy, et cetera.

In this project, the focus of the research is on general-purpose industrial research institutions. Even after we limit ourselves to this one type, we find that their potential goals and functions can still remain quite diverse. The functions usually encompass increasing scientific knowledge, producing newly trained scientists, undertaking a certain amount of regulatory and standards work, service functions for industry, and provision of contract research, developing and building prototypes and test equipment, and, finally, performing the roles of a private R&D lab within a firm whose main objectives are to reduce the costs of existing products, generate innovative products and processes, reduce the time lag with competitors, and develop new designs and processes that adapt to changing supply, price, and demand signals. It is clear from this list of objectives that for many national industrial research institutions, a possible problem may be that the list of goals and objectives is too diverse and too numerous for the human and financial resources available.

We find that, even within the innovation functions of the ITRI, we have to differentiate between different types of innovations that a particular lab may be engaged in. Marquis (1969) distinguishes between three different types of innovations and their characteristics. The first type is concerned with the management of technological change needed for very complex decisions, such as communication networks, where large sums of money are involved. This type of innovation is characterized by thorough, long-range planning that assures that the requisite technological subcomponents will be available and ready to be assembled at each stage. This type of innovation is specialized and not commonly demanded for most ITRIs.

The second type of innovation concerns breakthrough technologies that change the character of an industry — for example, the jet engine, xerography, or the oxygen converter. Such innovations are rare, unpredictable, and usually produced in large research labs, often outside of the sectors that are most affected by them. They usually come from outside because people within the industry tend to be preoccupied with short-term concerns and focus on improvement, quality control, cutting costs, expansion, and so on. The third type is the "nuts and bolts" innovation, which is essential to gaining economic benefits from technological change. These are more intimately paced by economic factors and involve continued improvement in product characteristics and process efficiencies. While we may hope for or demand innovations of the second type from **ITRIs**, they are not common, they cannot be programmed, and their impacts are usually very broad and take place over a long time — for instance, the development of the microchip — and so they are most difficult to plan for and to measure.

1.1 Evolution of Public ITRIs

To place the development of ITRIs within a historical context, we note that in OECD (Organization for Economic Co-operation and Development) countries, that public research naturally emerged in areas of dominant state interest. It started in the field of defence, the second phase was the setting up of geological surveys and agricultural research. Then later came industrial research institutions often originated to establish standards and maintain regulatory controls, not to generate industrial innovations. Examples include the Laboratory of the Government Chemist in the United Kingdom (1842), followed by the National Physical Laboratory (1900), the National Bureau of Standards in the United States (1901), the Physikalisch - Technische Reichsanstatt (now PTB) in Germany (1887), and the first such institution was established in Japan in 1913 as the Central Inspection Institute of Weights and Measures. It is only among the later industrializers and the smaller countries that a more deliberate attempt was made to use the industrial laboratories to develop and diffuse industrial technologies. Examples of this direction include the Technological Institutes of Denmark (1906) and the Danish Academy of Technical Sciences (ATV), with 19 research institutions in 1937; the National Research Council of Canada (NRC) (1916); the Australian Council on Science and Technology (1916); and the TNO in Netherlands (1932) (OECD 1987a). The developing country institutions were often designed in imitation of institutions in the United Kingdom, the United States, and France. Thus, while at one level their principal objectives remained pragmatic, as for the late industrializers, to develop and diffuse technologies to industry, their model was more often oriented towards basic research and in variance with their pragmatic objectives.

The post-World War II period was one of optimism and faith with respect to an industrial R&D revolution. Over the next three decades, there was a considerable and continued increase of resources for R&D, both in absolute terms and as a percentage of the growing national product. During these decades of prosperity, it was almost axiomatic that more science was a good thing and that increased expenditure for R&D was required to reap these benefits. This view was seemingly supported by research that showed that a substantial (although the shares are widely debated) portion of economic growth was accounted for by the increased scientific inputs and resultant technological change. The steady, rapid, and unprecedented economic growth rates of this period allowed more resources for all activities, including steady increases in public R&D spending without much debate and dissension. While many commented that public R&D spending could not keep growing at this rate forever, the day of reckoning always seemed far away. Over time and with the establishment of many and varied types of industrial research institutions in all countries, many of the national institutions moved larger fractions of their resources towards basic research and towards developing knowledge at the scientific frontier. However, with the resource scarcity of the recent decade, their inclinations came into increasing conflict with the users, requiring many of them to confront the questions of effectiveness, efficiency, and return on investment that confront the research institutions in Latin America.

In this context, a study by the World Association of Industrial and Technological Research Organizations (WAITRO 1972b) is illuminating. In this study, the directors of 49 national industrial research institutions in developing countries were asked to express the needs of their institutions in order of priority. The directors of the ITRIs placed the need for economic justification of their institutions within the national economic context at position 30, the lowest-priority need. Most of the topics discussed in this paper, such as evaluation, strategic planning, and organization structure, were in the bottom five of priorities identified. Their assessment of priorities probably reflects well the dominant view of research managers in the industrial countries at about the same time. However, over the past decade, the issues that were accorded low priority have risen to the top, and the large number of recent studies of the effectiveness of R&D that we review here is a testimony to the changed environment.

Notwithstanding the resource scarcity, or perhaps because of it, there emerged in the late 1970s an increased demand for relevance and accountability of government-owned research institutions, a demand that was accentuated in the 1980s by public sector deficits and an attitude essentially and almost ideologically hostile to the entire public sector (OECDa 1987, p. 71). These financial and political dimensions have been reinforced by a rapidly changing economic and technological environment that calls into question the original or historical roles and missions of these institutions.

We will examine later some of the assessments undertaken in the OECD countries, some of the methods used, and a general guide to new organizational patterns. Before that, we need to review the question of efficiency, in particular that of public research institutions. To do so, we need to determine the inputs and outputs of research institutions and how they can be measured, and then we need to see whether institutional forms make a difference in the inputs, outputs, or other significant intermediating variables. A general point that may be made here is that effectiveness requires a certain clarity of goals, objectives, and missions, and these have to be defined in relation to the resources available and the particular environmental context for the demand and supply of knowledge. We need to make sure, both in our assessment of the past performance of these institutions and while recommending new directions, that we take these contextual and historical factors into account.

2. INPUTS, OUTPUTS, AND MEDIATING FACTORS

2.1 Inputs

The critical inputs for a research institution are acknowledged to be its scientists and engineers, as well as the facilities and equipment available. However, it is also well recognized that the actual numbers of scientists and engineers alone have far less bearing on the output than their levels of creativity and other even "less adequately defined and still unmeasurable" (Gold 1989, p. 60) aspects of context. Some of these additional variables, including structures, procedures, organization, networks, and linkages, are discussed later. For the moment, if we can subsume these other aspects, we can attempt a measure of efficiency, with the denominator as the number of scientists, the total financial resources, or the resources available per scientist. The input side, to that extent, is, compared with the outputs, at least relatively conceptually simple and straightforward to calculate at a program or institutional level. We will see in the subsequent discussion that the transformation process of the inputs to effective outputs is highly complex and strongly mediated by multiple factors, some within the system boundaries and others external to it.

The accuracy of input estimates has been found by many authors to be variable and especially inconsistent in the industrial sector. Studies by Meadows (1968), Allen and November (1969), and Allen and Norris (1970) show a general underestimation of costs, and Mansfield (1968) found initial R&D cost estimates to be quite inaccurate.

2.2 Outputs

The measurement of the output side of a research institution is fraught with many more problems. The outputs of the R&D process or of an institution consist very broadly of new knowledge pertaining to specific problems. However, this characterization does not take us very far in determining the quantity or quality of the outputs. The new knowledge generated by the researcher and the institution must be communicated to others, used by others, or have some influence on other actors and agents to be socially meaningful. That provides us with a way of identifying and listing the different outputs that can be relevant.

As all knowledge must be communicated to be socially useful, the first and most important group of outputs consists of the means used for the communication of new knowledge. These may take the form of internal or external reports of limited circulation, publication in peer-reviewed journals, presentation of results at conferences and seminars, publication of books, or contribution to public awareness of and participation on the issue. Other written outputs may include the development of course materials, training programs, standards, and specifications. If the training of new researchers is judged to be important, then oral and written communications with students, trainees, and young researchers have to be added on the output side.

For example, the National Research Council of Canada (NRC 1992) lists its objective and outputs for engineering research as follows:

Objective:

PREFACE

This paper originated as a contribution to a research project on the changing role of industrial technological research institutions (ITRIs) in Latin America. The research project, consisting of a number of case studies as well as this background paper, has been supported by the International Development Research Centre, Canada. The key objective of the project was to examine directions and approaches towards making the Latin American ITRIs more "effective."

This background paper was initially prepared to assist the researchers in undertaking the Latin American case studies and determining appropriate recommendations. Its objective is to contribute to the research studies through a discussion of the concepts of the "effectiveness" of research and development (R&D) and to cull from the literature methods that have been suggested or used to make R&D institutions more "effective." A shorter draft was presented at a meeting of the researchers in Caracas in 1993.

Based on the comments and feedback received at the Caracas meeting and subsequently, we concluded that a more detailed discussion of the issues could be useful to other researchers studying R&D institutions and also perhaps to managers of such institutions. We will be pleased if this review of the literature serves any of these wider purposes.

Amitav Rath
Ottawa, April 1994

To undertake and promote engineering research and development in support of Canadian economic development in strategic areas of national importance, such as transportation, resources, construction, manufacturing, and information technology.

Outputs:

Technological procedures, processes and products;
Reference materials and product standards;
Technological information;
Industrial research support and collaboration;
University support and collaboration;
Research and facility services;
Training in research methods and procedures;
National engineering facilities usage;
Research contracts, revenue and patents; and
Trained engineers and researchers.

While the quantity of these outputs can be determined by listing and counting, determining the quality of outputs requires more work. The qualities of the research outputs are usually judged with regard to their effectiveness, efficiency, quality, profitability, and degree of innovativeness (Armentrout 1986), and we discuss these issues in the evaluation section.

2.3 Environmental Context

In much of the discussion of R&D institutions, as with other institutions, it is considered that the differences between public and private research institutions are important. It is often further assumed that private sector institutions are more efficient, support greater application, and are hence more effective.

Studies by Pfetter and Long (1977) and others suggest that organizations that are more independent tend to be more powerful and more effective. The budget allocation process, civil service rules and systems, public procurement and contracting systems, and financial management and audit procedures, all externally determined, tend to increase external constraints on public sector ITRI in comparison with private sector ITRI (Bozeman and Loveless 1987). Cordell and Gilmour (1976) suggest that public and private research laboratories respond to different sets of stimuli.

Government organizations are often seen to be more bureaucratic and to suffer from excessive formal procedures. However, one study by Pelz and Andrews (1966) found little difference between large government and private industrial research organizations in terms of procedures.

There are a number of studies on organizational climate and motivation that invariably suggest that participatory structures and decision making, high degree of job satisfaction, and motivation lead to more productive research staff. Rainey (1979) suggests that there is a weaker link between performance and recognition in the public sector, thus reducing motivation.

Bozeman and Loveless (1987) ask the question as to whether the differences between the public and private sector ITRIs are due to the effect of ownership or to different mandates, different sectoral emphasis, or different types of activities undertaken. They go on to suggest that public and private sector ITRIs face two different types of constraints, one the market, and the other political and administrative constraints, which in turn lead to different organizational climates and motivations. In the pure form, a private organization provides goods and services that are exclusionary, are appropriable, and maximize profits. A public organization, on the other hand, is sheltered from the market and would tend to produce information and innovations on a nonexclusionary basis, which may in turn be privatized by other agents. It should follow that, in a public ITRI, publications in scientific journals may be a more desirable output, whereas in a private ITRI a patent or new product may be more valued, thus leading to different types of outputs.

Bozeman and Loveless (1987) investigate a group of 50 U.S. labs — all focused exclusively on the engineering sector, thereby eliminating sectoral variations — to answer the question as to whether the attributes and behaviours of research units such as the work climate and red tape are a function of "public-ness" and "private-ness." Their findings support the hypothesis that the influence of external control and oversight reduces the level of autonomy of the research institution. They find that the market-based constraints on private ITRIs work through the organizations' own leadership and tend to work more as signals than as commands. They agree with the report by Buchanan (1974) that managers of public sector ITRIs have less authority, and to that extent public-ness should affect creativity and innovativeness negatively.

However, Bozeman and Loveless (1987, p. 214) report that, contrary to expectations, the public units in their sample had lower work pressures, indicating a more relaxed and less intense work environment, a higher morale, and a willingness to accept new ideas, all leading to a more positive work climate. They found support for other elements of conventional wisdom: reward structures in public ITRIs were less well organized; and, while public labs had somewhat more public outputs such as books, reports, and papers, they had fewer appropriable outputs such as patents and internal reports. In their sample, outside evaluators rated the outputs of the public laboratories to be "slightly more valuable to society." They caution against an uncritical acceptance of the "popular mythology (and some case evidence) that the performance in the public sector is not up to the standard of private sector."

Yet the popular mythology survives, perhaps because there are at least sufficient anecdotal and personal experiences of nonperforming public institutions. In developing countries, which have poorer resources, a shorter institutional history of ITRIs, and weaker articulation with demand, the pathologies can become large and grotesque. Consider the following observations about an institution in a poor developing country: "Inter and intra bureaucratic conflicts may lead to institutional 'immobilism'...The situation is the more alarming because in terms of manpower these bureaucracies grow day by day since the State is not only committed to training high level manpower but to providing jobs as well in these bureaucracies...This problem acquires pathological proportions in a country with limited resources...Parasitism is evidenced through misuse of public funds, overemployment, and the heavy government subsidies" (Mudola 1983).

In many developing countries, the public sector is often plagued by deteriorating economic conditions, poor governance, overcentralization, lack of incentives for civil servants, and a brain drain. Organizational restructuring and privatization have been the most frequently suggested means of reform, but difficult questions remain as to how far the public sector institutions can be strengthened without addressing wider reform issues, such as system of governance, administrative culture, and the structure of the civil service (Plowright and Wills 1994). Plowright and Wills (1994) summarize the difficulties often faced by public institutions in developing countries:

Country Environment:

1. Lack of government commitment,
2. Turbulent conditions (social and political changes; frequent change in objectives and priorities),
3. Restrictive policies and procedures (import policies, bureaucratic bottlenecks, etc.);
4. Lack of qualified indigenous staff.

Weaknesses in Internal Management:

5. Unrealistic goals or poorly defined objectives,
6. Unrealistic time frames,
7. Weak motivation and leadership,
8. Short-term expediency over long-term capacity building.

Poor Institutional Arrangements:

9. Unclear lines of authority,
10. Excessive coordination and linkages,
11. Parallel implementation.

External Effects:

12. High dependency on external donors,
13. Dominance of foreign experts.

Poor Fit in the Local Environment:

14. Insufficient account of local conditions,
15. Inappropriate technology,
16. Failure to involve intended beneficiaries.

2.4 Motivating for Creativity and Innovation

People are central to any research institution. One view is that what differentiates the star research performers from the good are inborn traits. This is disputed by Kelley and Caplan (1993), who report initiative, networking, self-management, teamwork, leadership, followership, perspective, show-and-tell, and organizational savvy as the key ingredients for better performance. They also found that "meetings and more meetings" were the largest-cited obstacle to productivity. To improve on performance, the scientists and their managers can be trained on these issues. In a review of some main psychological and organizational theories, Angle (1989) points out that an industrial research organization must motivate creative people to join, to stay, and to perform both reliably and in innovative ways. Motivation and ability of people go hand in hand, and they are both related to the individual and to his or her environment.

Angle suggests that more individualized reward systems stimulate creativity and radical innovations. Group-based reward systems are more effective for implementation programs and for incremental innovation. For either type of reward to be effective, a performance appraisal system is needed that is open, transparent, fair, and equitable — and, more important, seen by the researchers to be so. In fact, Kelley and Caplan (1993) found in a study of the performance ratings of scientists at Bell Labs that there was only a 50% agreement between the ratings given by peers and those given by management. As we discuss subsequently, with the many difficulties of evaluating the quantity, quality, and effectiveness of research designing a suitable system is not simple (Beer 1988), but at the same time the value of simple and effective evaluation systems for motivation cannot be overestimated.

In reviewing the supportive organizational factors, Angle (1989, pp. 164-165) lists a number of propositions on innovation. He states that even highly motivated and capable researchers find it difficult to be creative when they lack access to basic resources of money, information, materials, and time. However, there are a number of enabling actions that can increase organizational innovativeness. These include high frequency of communication and a moderate, not large, amount of environmental uncertainty. The level of innovative activity increases with the existence of mechanisms to focus attention on changing conditions. Bowman (1992) develops a group of indicators to measure the role of the research manager in generating better outputs.

2.5 Problems and Success

Our outputs need to be reviewed in terms of their value. In an overall list of key problems often associated with research institutions, which the ITRI manager must intervene to eliminate, Abend (1985) suggests that some typical problems include:

- solving the wrong problem,
- solving yesterday's problems,
- fossilized thinking,

- N.I.H. (not invented here),
- user needs never analyzed or specified,
- insufficient alternatives examined,
- limited imagination in planning or design,
- fragmented development (R&D or engineering),
- high frequency of program aborts,
- diluted programs,
- efforts begun too late,
- organizational inertia or resistance,
- hostility or biases across group lines,
- overdomination by a single group,
- expected results don't match effort applied.

Mansfield and Wagner (1975) suggest that R&D is an activity involving three sources of uncertainty — technological, commercial, and economic — and so the final results need to be judged along each of them, although we may ultimately be interested only in the final metric. Baker et al. (1986) reviewed the technological and commercial dimensions of project success for 211 R&D projects from 21 companies in four fields and found a 50% success rate at the technological and commercial levels. This study showed that project success was more likely with earlier experience in production, the market, and/or the science and technology (S&T) area. Surprisingly, well-defined business and technological goals at the *initiation* stage were not significantly related to eventual success. Those projects expected to result in exclusively new or modified processes were more likely to succeed (69%) than those expected to result in new or modified products (48%). At the same time, we should note that some studies show that the latter have a higher economic payoff associated with success than the former. Projects begun at the suggestion of marketing/distribution/sales and/or customers were more successful (56%) than when R&D was the sole instigator (35%).

Marquis (1969) examined more than 500 innovations in products or processes that occurred in 121 companies over the 5-10 years prior to his study. The innovations were judged by executives as the most important to their firm. Upon review, it is evident that innovation is a total process of interrelated subprocesses, not just the conception of a new idea, the invention of a new device, or the development of a new market. Innovation may be carried out from conception to implementation within one organization, but more commonly contributions come from other sources at different times and places. Successful innovation begins with a new idea, which involves the recognition of technological feasibility and market demand. Both are essential. Idea formulation is the fusion of a recognized demand and technological feasibility into a design concept. Part of this stage is in fact evaluation, which will recur along with the process of innovation. However, a strong judgemental input is needed here in order to secure a commitment to resources.

Next comes problem solving. More information may be necessary, and unanticipated problems may arise, calling for new solutions or trade-offs. Great obstacles often force the termination of the project. If a solution is found, the technological feasibility of the original demand or a modified one has been demonstrated. Development comes next. Innovation is never achieved until the product or process is marketed and sales or cost reductions are achieved. The solution is first utilized and diffused in the marketplace. Only one or two products out of five break even on their investment.

What factors are involved in successful nuts-and-bolts innovations? Marquis (1969) found that (i) small, incremental innovations contribute significantly to commercial success, (ii) recognition of demand is a more frequent factor in successful innovation than recognition of technological potential, and (iii) the training and experience of the people are the principal sources of information for successful innovations — effective selection, development, retention, and use of personnel are management's top priority.

The need for involvement by marketing departments as opposed to the need for good market assessment receives the least attention by researchers studying research. Roberts (1959) points out the need to involve marketing researchers in product development but does not define when they should become involved. That they should be involved early in the idea evaluation is stressed by Treeger (1969). Kegerreis (1969), Cox (1972), and others consider early involvement as critical for success. While the need for the involvement of marketing researchers and users from the initial stages for ensuring higher rates of success is supported by a number of studies and is the driving force for many of the reorganizations of ITRIs around the world, there are some who have argued against it. Bedrosian (1971) argues from his experience as the manager of R&D that "Marketing should be brought in before an R&D project is initiated, rather than at the test-marketing phase, so that no time is lost persuading Marketing of the product's merits. Their input at the proposal stage will enable one to build into a product the features which will make it a success. But, if Marketing dominates the R&D lab, it might seem to have more realistic objectives but its products will be imitations of what already exists." Thus, some uncertainty about *exactly when and how* marketing and R&D should collaborate is not resolved by the literature.

Ali et al. (1993) demonstrate that the potential uncertainty involved in developing and commercializing new products inhibits firms from committing to truly innovative projects. Such projects have been shown, though, to provide handsome returns if commercialized. Optimal project selection requires a trade-off between the size of the returns expected and the risk associated with them. Risk comes from technological uncertainty (uncertainty in project completion time given a certain development cost function) and market uncertainty.

3. EVALUATION

Evaluation of research results remains one of the most demanding and subtle tasks (Wallmark et al. 1988). However, in the light of scarce resources, evaluations of scientific activity have become institutionalized in all major countries carrying out research. They assess the effectiveness of meeting social, economic, and intrascientific goals and the development of infrastructure for training and research program support. Science must be evaluated in terms of what society expects from it, based on autonomous goals of science or its larger utility. The importance of defining criteria has been debated since Weinberg's (1964) classic paper on intrinsic and extrinsic criteria. Evaluation should be seen as an analytical and interpretive process that seeks to establish interrelationships between outcomes and the contextual preconditions of scientific performance (such as material, human, organization, cognitive resources) upon which to base policy. It also calls attention to the need to distinguish between results upon which evaluations are based, their collection and systematization, and their comprehensive interpretation. Conceptual distinction is necessary between indicators, monitoring, and evaluation of performance (Stolte-Heiskanen 1986).

We must first distinguish between the many different types of evaluations, their purposes, and therefore the necessarily different methods that are applicable. Evaluations are required, at various points in time, to select the individual research projects to be undertaken and to monitor and decide on the continuation of ongoing research projects. Finally, they are required at or after the termination of the research project to determine potential exploitation and to determine what was achieved (Christensen 1987). Some time after the project termination, evaluations can serve to provide an assessment of the quality and the impacts of the outputs produced. Is quality measurable in terms of Nobel prizes or are such events unique events? Can investments in science be compared with other types of investments? Are differing styles of research support better for different fields of science? We review some of these issues further. The impacts may be on the scientific community when measures such as citation counts can be used as an indicator of quality. Or the impacts may be in the economic domain, in which case the methods of benefit allocation and the time of measurement become crucial issues.

Evaluations may also be differentiated by their purpose: to raise an issue, formulate new policies or programs, define alternatives, improve existing programs, mobilize support, change ways of thinking, or plan new research (Chubin 1987). It is important to distinguish between pre-project, ongoing, post-project, and impact evaluations (Collier and Gee 1973). The proper timing of evaluations is important. Should they be done as an input to a specific decision or as part of the agency/program archives? The rationale for evaluations and their design must reflect the needs of both decision makers and research communities.

3.1 Project Selection

Christensen (1987) suggests that often in practice the maximum effort is put in at the pre-project stage into the process of selection of the research projects to be undertaken. The next phase of ongoing monitoring tends to be informal and receives less attention than the initial procedure to accept or reject. In the past, the final post-project evaluation often received the least importance. A survey of Industrial Research Institute members confirms the importance of selection process and the importance of the initial decision in the productivity of R&D. Mansfield and Brandenburg (1966), in a study of a central research laboratory of a U.S. firm, conclude that estimates of the probability of a project's technological success made before a project is started have some, but not much, predictive value of final outcome. Yet, as we will see subsequently, the traditional distribution of emphasis, especially common to public ITRIs, is not an optimal use of resources. It is more often based on procedural propriety, that is to satisfy audit requirements for expenditures of funds. It is at the initial stages, before any resource commitments are made, that the evaluation and selection has been the most rigorous in public research institutions.

3.2 Mathematical Models for Project Selection

Convinced of the importance of useful methods for selection of individual R&D projects and encouraged by the success of various mathematical programming methods for optimal decisions and portfolio selection models for financial assets, many management scientists have proposed different types of models to assist the selection process. Examples include Asher (1962), Cramer and Smith (1964), Dean (1967), DeCicco (1968), Lucas (1971), Gear et al. (1979), and many others. There has been a major attempt by management science researchers to model the processes for R&D project selection, and generally they attempt the maximization of expected utility and utility maximization principles were first incorporated into mathematical programming models in the late 1960s. (For more examples, see Booker and Bryson 1985.)

The models can be broadly grouped into two categories: benefit measurement and project selection/resource allocation. The former type all require well-informed experts to provide subjective input for project proposals and integrate objective benefit data. Among the benefit measurement methods for selection there are a number of approaches. The **first category, comparative approaches,** uses approaches like Q-sort, ranking, rating, paired comparisons, dollar metric, standard gamble, and successive comparisons. **Scoring models** use a small number of decision criteria such as cost, manpower availability, scheduling feasibility, and probability of technological success. The resulting vector of scores for an alternative is useful as a diagnostic for selection. The project scores are then sometimes combined, usually by addition or multiplication, to get an overall benefit measure, or they may be left as a vector of scores. This approach has received increased emphasis since 1969. **Benefit contribution models** require that projects be tied directly to R&D objectives or to systems requirements, not specific project characteristics. The numerous economic return, cost/benefit, risk analysis, relevance tree, and assessment tree approaches are examples. Benefit is measured in terms of contributions to a number of objectives or systems. The selection models tend to use mathematical resource allocation techniques.

Baker and Pound (1964) cite 80 selection methods for research projects from the literature but found no evidence that managers were using many of the methods or models of selection that had been proposed. A 1973-74 survey of Industrial Research Institute members revealed that only one-third of the 193 respondents used any formal project selection methods. Formal methods in use were designed primarily for collecting information along different dimensions for new or ongoing projects but did not imply the use of complicated mathematical or decision models reported in the literature.

Brandenburg (1966) reports on an empirical review of ongoing project decisions at 14 companies and found a common pattern. Adding and deleting projects, reprogramming active projects, replacing ongoing projects with modified ones, and shifting the resources among projects were common ongoing managerial changes. In addition to the initial selection of projects, the selection process also includes generating new alternatives, determining the appropriate times for decisions, collecting data, and specifying constraints and criteria. Brandenburg found that profit maximization is not a sufficient criterion and accounted for only 50% of allocation variations. Predictability of results and the likelihood of technological success; the perceived future effect on sales volume or revenue and on savings in material, labour, and other costs; effect on profits; time and cost characteristics of the project; and customer satisfaction were all simultaneously factors of concern. He also found that a progressive change in criteria for filtering new project proposals from scientific merits to economic potential occurred as projects moved from the "R" towards the "D" of the R&D spectrum.

Accurate numerical estimates of contribution of each project to individual criteria cannot be easily generated. Difficulties abound for estimating benefits owing to the multiple criteria, which lack a general underlying measure and whose relative importance varies over time. In essence, the initial selection process is not that of a simple "go" and "no go" but is really part of a continuum where selection merges into monitoring. These observations led Connolly (1972) to suggest that *all models of project selection were inadequate in their assumption of a single decision maker and that they do not incorporate the extended time frame for research and the importance of structures and linkages. These ongoing processes are diffuse in nature, lack definite end-points, have unclear alternatives and uncertain cost-benefit estimates, and have conflicting preferences with changeable resources and constraints.* In a later review, Baker (1974) characterizes most papers on project selection to be speculative and simplistic.

Baker and Pound (1964), Cetron et al. (1967), and Baker (1974) provide some of the early surveys of the literature on the various quantitative methods for project selection. These reviews all suggest common limitations that are inherent to the abstract models proposed:

- inadequate treatment of risk and uncertainty; of multiple, often interrelated, criteria; of project interrelationships with respect to both value contribution and resource allocation,
- no explicit recognition and incorporation of the experience and knowledge of the R&D manager,
- the inability to recognize and treat nonmonetary aspects such as establishing and maintaining portfolio balance and linkages,
- inadequate treatment of the time variant property of data and criteria and the associated problem of consistency in the research program.

Given these limitations, "the common approach...to quantify preferences or subjective estimates of benefit with unsatisfactory methods," and the perceptions of R&D managers that the models are unnecessarily difficult to use and understand, it is no wonder that the models are not commonly applied. The Baker (1964) review demonstrated a lack of implementation and identified the need for formal, empirical studies of the implementation process. Several cases are cited in which resource allocation models were adopted but then discontinued when the model builder or sponsor left the R&D organization. Given this, applications started moving away from "decision models" and towards "decision information systems."

Yet later Schroeder (1971) finds that the number and sophistication of models designed and proposed for project selection kept growing exponentially in tandem with the growth of new computational technologies, despite the infrequency with which they were applied. He believes the most promising models are those that (i) focus on evaluation rather than the decision, (ii) indicate preferred rather than optimal courses of action, and (iii) are capable of dealing with qualitative elements. Gerloff (1973) again found that although a great number and variety of techniques have been suggested in the past decade to improve the technological, cost, and schedule performance of government R&D projects, the difficulties persisted. He tried to determine if the techniques were helpful, using performance and control data from 108 government-sponsored R&D projects between 1950 and 1967. His results showed that the projects having a large volume and variety of control techniques applied could not be statistically associated with performance improvement compared with projects not having such control. In fact, he found that rather excessive formalization was linked with greater failure! Such findings led Collier and Gee (1973) to comment that "much has been written about the necessity and difficulties of research evaluation. It is difficult to know objectively whether the suggested methods for the improvements in project selection, creativity, communication, etc. are effective." They go on to conclude that there is probably no organization with the answer to the measurement of research effectiveness, and certainly no generally accepted principles for post-evaluation of R&D. The lack of utility and applications did not prove to be a deterrent to the further development of the techniques of project selection. Clarke (1981) provides over 3000 references to decision models for R&D management, and their lack of use is again discussed by Cardus et al. (1982).

Booker and Bryson (1985) provide a more recent overview of the subject of mathematical models and decision theory for research managers. They point out that "gaining the required knowledge of decision theory and decision analysis (which should be very useful in the art of good project management) is difficult." They argue that decision-theoretic optimization assumes that the best decision maximizes utility, but it is difficult to determine the relevant utility functions, and a major problem is the amount of quantified information required. Some models optimize conventional attributes such as expected profits, discounted cash flow, or time constraints like budget, skills available, facilities available, raw materials available, risk, and overall program balance. Other model approaches have included dynamic programming methods with stochastically (varying in time) characterization of the probability of project success, and goal programming approaches. Other mathematical programming techniques, especially integer programming, were developed to create a portfolio of accepted projects whereby projects are included or excluded depending on resource allocation and other constraints.

Mathematical models are complex by nature (their greatest drawback) because of the need to include all features of reality such as constraints, cost functions, resource bounds, utility functions, and weighing factors. Booker and Bryson discuss alternative schemes, including averaging ratings over several experts, partial preference ordering, and graphical techniques. In reviewing the problems of using these tools, another author demonstrated variation in results using the same model within the same company, which only adds to the reasons why so few models for project selection have been attempted. Krawiec (1984) agrees that mathematical models have not been used much despite their popularity in the literature and suggests that simple scoring methods can incorporate many of the relevant variables and are a greater help in deciding. Doz et al. (1985) discuss the fact that the success or failure of one project often depends on the results of others because of interdependence.

Human behaviour adds another complicating dimension for decision models. People tend to overestimate punishment and confidence limits and underestimate rewards and time. Small probabilities tend to be overweighted. Risk taking is more common when dealing with loss than with gain. Studies show that model data that are often derived from expert opinion are still plagued with biases. Estimates vary randomly around the true (but unknown) value being sought. Experts tend to underestimate probabilities of variation and overestimate the accuracy of their own judgement. The modelling techniques from decision theory and mathematical programming tend to restrict the users to rigid formulations, thus losing accuracy. More importantly, flexible techniques such as scoring and ranking and statistical methods keep the user involved and thereby improve understanding of the process.

Sanchez (1989) notes that selection models have mostly been produced in the 1950s, 1960s, and mid-1970s. Since then, the criticisms noted above have deterred the development of many more, although examples such as English and Czerwon (1990) suggest that some continue to have faith in mathematical approaches. Newer philosophies for evaluation and selection methodologies have been developed that emphasize the decision processes, human factors, and preference structures. Sanchez classifies four groups of strategies based on a sample of innovative companies in Spain:

1. a planning strategy,

2. an economic strategy with multiple criteria set by top management for evaluation,
3. a market strategy with input from R&D, and
4. a technological strategy where the technological element of innovation is critical.

Companies with planning strategies use various *ad hoc* methods for selection in order to achieve company consensus. For companies using economic strategies, criteria tend to be fixed and selection methods are inflexible. Evaluation methods are used solely to rank projects for selection for companies with a market strategy. In the last group, technological strategy, the highest weight is given to innovative dimensions.

Having reviewed the various methods and their advantages, we conclude that in practice all methods of project selection tend to assess research projects regarding their potential contribution to a number of different objectives. A checklist of commonly used criteria for making judgements would often include the following:

Corporate objectives:

- fits into overall objectives and strategy,
- corporate image.

Marketing and distribution:

- size of potential market,
- capability to market product,
- market trend and growth,

- customer acceptance,
- relationship with existing markets,
- market share,
- market risk during development period,
- pricing trend, proprietary problem, geographical extent, and effect on existing products,
- completion of product line,
- quality improvement,
- timing of introduction of new product,
- expected product sales life.

Manufacturing.

- cost savings,
- capability of manufacturing product,
- facility and equipment requirements,
- availability of raw material,
- manufacturing safety.

R&D:

- likelihood of technological success,
- cost,
- development time,
- capability of available skills,
- availability of R&D resources,
- availability of R&D facilities,

- patent status,
- compatibility with other projects.

Regulatory and legal factors:

- potential product liability,
- regulatory clearance.

Financial:

- profitability,
- capital investment required,
- annual (or unit) cost,
- rate of return on investment,
- unit price,
- payback period,
- utilization of assets, cost trend, cost reduction, and cash flow.

Becker (1980) cautions that such a list must be generated for each organization, keeping its unique characteristics in mind. He finds that generalized checklists for project selection have limited value owing to differences in business objectives and strategies. In one industrial lab, all information on research, profitability goals, and success probability was brought down to a single page (Bedrosian 1971), presumably because many of the items in the checklist above had been subsumed into the larger planning processes.

3.3 Post-Project, Output, and Impact Evaluation

The need for evaluating the effectiveness of an institution arises from many sources, which have been noted. Evaluation research is an industry that largely emerged from education and social psychology and applies a range of techniques to satisfy its need to evalurate. While some of the findings of the general evaluation literature are pertinent to our purpose, most of them are not. The model of evaluation that has the highest scientific validity is the "Quasi-Experimental Design" (QED). Here, researchers attempt to apply the experimental model as practised in social psychology, comparing differences between affected and unaffected groups, with the help of "representative" samples (Plowright and Wills 1994).

This approach presents problems, especially with research evaluations. Research designs for such evaluations are costly and elaborate. Other problems include an overemphasis on quantification and measurable "impacts." Furthermore, the QED model requires clear objectives, but researchers in the field usually find a ~~tangled yarn~~ of competing goals. QED also often means that local contexts are neglected in order to yield generalizations about cause and effect. We will see that many of these contextual factors are major determinants of failure or success. Traditionally, an "evaluation" often meant a one-time descent of "objective" external experts. Often, their theoretical model was irrelevant to project managers, and so uses of the evaluation findings were few. This reliance on external teams meant lost opportunities to utilize internal capabilities; increased risk of sociocultural-political contextual errors; and greater likelihood that findings would not be used.

In the 1980s, dissenting voices argued that the use value of an evaluation was more important than scientific rigour. More realistic methodologies began to use small samples, open-ended interviews, and quick proxy impact indicators. The notion of "impacts" broadened to include various side effects, and multiple goals began to be accepted. Modified approaches are more subject and context specific and ask what is the research about and what are its goals and its anticipated applications. Currently, the long-term accountability of public funds for research is measured in terms of quality of knowledge and scientists produced. Accountability can be viewed at the financial level — i.e., whether the funds are properly spent — and at the project level, which examines its technological success.

A controversial evaluation approach that involves "converging partial indicators" was developed recently by Irvine and Martin at the Science Policy Research Unit, University of Sussex. Despite such problems as collecting, comparing, checking, and reconciling data, they argue that the approach offers a comparative methodology for systematic assessment based on returns on investment. Many criticisms have been made of the method, because administrators have used it to cut vulnerable projects.

A subject that is not often raised in the literature of evaluations, which tends to have a rationalistic bias, is the intrusion of political processes into any measurement, evaluation, and resultant judgements. These influences are likely to be higher within public ITRIs and in the evaluation of public science. Some of these issues are discussed by de la Mothe (1992), and others have been noted earlier in our analysis of the changed environment for public ITRIs and subsequently in the case studies discussed by Oldham (1988) and Araoz (1993).

Separation of the evaluation processes of R&D into a multistep process allows for the appropriate distinctions between the needs of project selection or resource allocation, monitoring, and results evaluation. Only then can the time spans involved for the different purposes be set, providing useful feedback to management. They can indicate whether problems exist and whether the research portfolio is in balance. Appropriately defined evaluation provides a common language for researchers and management and encourages communication between research and its sponsors and users.

The multiple steps to determine success or failure necessarily involve the separation of the technological performance of the research staff to develop solutions related to the objectives from the economic value, both potential and realized. Then only comes evaluation of the potential value of what the research staff has produced for projects whose responsibility has been transferred. This answers how well is research contributing to commercial objectives? Low productivity could arise either from the failure of researchers to achieve the specific technological tasks agreed upon with management or it could arise if the completed tasks do not contribute to commercial objectives.

A simple rating scheme is suggested by Collier and Gee (1973) to measure performance. Objectives should be stated in terms of definable performance that can be understood and measured. A system to rate technological success is suggested, involving a four-point scale of 0 through 3. A 0 would be assigned to a project that has missed its objectives or that overran targeted costs or time; a 1 for projects with encouraging progress but where agreed-upon objectives are not met owing to changes in ground rules because of factors outside the control of the research project; a 2 for projects meeting objectives; and a 3 for projects exceeding objectives because of their technological success, because they were achieved under budget or sooner than expected, or because they went beyond the stage for the same amount of money. The rating could be further weighted according to the size of the individual project — i.e., be multiplied by the research expenditure or normalized over all projects (Collier and Gee 1973).

A number of useful overviews of the methods of evaluation are available. An early review of some of the issues is available in Glass (1970). Gibbons (1984) provides a good review of the literature in the 1980s and this is followed by OECD (1987b) which provides a succinct description of evaluation methods. Gold (1989) highlights some of the key problems in the evaluation of industrial research. A more recent review that focuses more broadly on institutional evaluations is by Plowright and Wills (1994).

There has also been an increasing trend towards multiple stakeholder participation in evaluations and the inclusion of ITRI management in the evaluation process. In general, the trend has been towards an acknowledgement of multiple goals; a movement away from rigid quantification; increasing attention to context; and emphasis on ongoing evaluations involving management. Perhaps most importantly, evaluations have become less of final "judgements" and more as "lessons" from which a wide audience can learn. Today, evaluations are considered most useful to increase the "capacity of learning" that must be built into all institutions from the start.

To summarize, let us assume for a moment that the industrial research laboratories only produce innovations that reduce costs, generate a new product, or substitute a less expensive input for another. In each case in principle, we can use the cost-benefit approach to estimate the set of benefits accruing from the innovation. Freeman (1969) cites two cases of such studies where the benefits were carefully calculated but warns that while this may be possible with some innovations, for others in general it will be too difficult.

If all innovations made are licensed and sold in the market, then the market value of the transactions can provide a proxy for the economic value of the output. However, for many public institutions the practice has been not to enter into market transactions. Where they have entered into market transactions, these have represented only a small part of their output. Further, with poorly developed markets, either outputs may be subsidized or excessive payments may be taken from captive customers. All valuation exercises are also plagued by the question of time. The value of an innovation can be properly estimated only after the lapse of a sufficient period of time for full development and impact to be measured. The time required for this is often too long to be useful for feedback into the planning process. Finally, inaccuracy of the profit measure, using conservative estimates at the planning stage, has often resulted in the aversion of long-term projects.

Other outputs of an industrial research institution include new findings, advice and consultation, training, testing, and patents taken out for new processes. New findings and some advice are provided in the form of scientific papers, technological reports, and presentations at conferences and seminars. To get a full measure of the outputs of research, we must make a count of each of these different outputs. Various problems arise in combining these different outputs into a summative indicator of the total output of the institution. While some of them can be converted to an economic value, many cannot.

The common approach is to first record the output in each dimension and then apply a qualitative evaluation of the significance of the papers, patents, and so on. Bowers and Elliott (1992) list the following evaluation criteria commonly used in the research laboratory:

- importance of problem being investigated,
- technological excellence,
- capabilities of investigators,
- adequacy of facilities,
- publications,
- leveraging of funds,
- number of students involved,
- patents,
- new discoveries or developments,
- significance of field,
- (potential for) technology transfer,
- importance of field to economic base of state,
- development of infrastructure for research and manpower development.

In applying these methods to estimate the quantity, quality, and significance of the outputs, a number of problems must be kept in mind. As with project selection, the next problem emerges — that of ranking the various criteria. Chan (1978) discusses some of the issues regarding ranking the different types of outputs, and some problems are discussed by Martin et al. (1987). We provide in the following chart a summary of some of the difficulties and their recommendations to minimize problems. For a more detailed discussion of these issues, especially from the perspective of small countries, Montgomery and Hemling (1987), Lukkonen-Gronow (1987), and Persson (1987), all of whom discuss the Nordic experience of evaluation, are useful.

3.4 Trends

In many cases, the problems of ITRIs are beyond the control of project management, although organization and management theories focus on structures and processes within the boundaries controlled by an institution. Such theories assume that the project units control all key resources and are virtually autonomous within their environments. However, success is determined by the project's influence over other entities, suppliers, ministries, beneficiary organizations, and so on. Smith et al. (1980) and Smith (1992) propose a new framework that deliberately looks outside the organization involved to include other stakeholders.

This moves away from typical "expert evaluations" that judge what is going right or wrong, but rather emphasizes a process that works with stakeholders to negotiate solutions to problems. This approach provides greater attention to stakeholders, including the beneficiaries. Such newer forms set out to be radical alternatives to previous types of evaluation. At the extreme, this approach abandons evaluation as a measurement-oriented, description-oriented, or judgement-oriented activity, moving to a new level whose key dynamic is negotiation. The paradigm holds that "evaluations" are not in any way descriptions of the "way things really are," but are only constructions that individuals form to "make sense" out of a physical, psychological, social, and cultural context (a " surround"). No constructions are "true" in any sense. Rather, they are shaped by the values of the constructors. As every society is value pluralistic, the question then arises as to *whose* values dominate. Evaluations can be shaped to enfranchise or disenfranchise stakeholding groups in a variety of ways. This approach particularly criticizes the "cosy relationship" that arises between evaluator and client, which gives preference to the objectives of the client over those of the stakeholders and tends to reveal weakness in groups *other* than the client (see Guba and Lincoln 1989).

Guba and Lincoln (1989) is especially critical of "context stripping" — that is, assessing the evaluand as though it did not exist in a context but only under the carefully controlled conditions. Such conditions are instituted in the hope that irrelevant local factors can be swept aside and more generalizable results obtained. Not only are such generalizations not possible, but surely this effort to derive general truths through context stripping is one of the reasons why evaluations are so often found to be irrelevant at the local level, leading to the much lamented nonuse of evaluation findings about which we, as a profession, seem so fond of complaining.

The newer orientation to evaluation is meant to define a course to follow and stimulate stakeholders to follow that course. Old-style evaluation relegated the task of follow-up to others; evaluators were then incensed when no follow-up occurred. Very often, however, the evaluation "product" was a set of recommendations that suited only the purposes of the evaluator and the client, with little concern for other stakeholders. If there is to be a course of action with which most stakeholders can agree, it can be arrived at only through *negotiation* that also honours the separate values of other stakeholders. The evaluators must play a larger role than simple information gatherer; they must orchestrate the negotiation process (itself the "guts" of the evaluation). For this, stakeholders are welcomed as equal partners in every aspect of design, implementation, interpretation, and resulting action. The idea is to work towards a consensual, more sophisticated joint construction. In fact, stakeholder input is meant to *determine* what information is collected.

One of the major tasks is to get each stakeholder group to deal with and confront the "constructions" of others. Through this process, the constructions of each group become more sophisticated. The evaluators then prepare an agenda for negotiation, and representatives from all the stakeholder groups join in, to arrive at conclusions. Those items that cannot be resolved remain as points of contention, but at least each of the stakeholders understands what the conflict is.

The proponents acknowledge many difficulties. They expect that the level of ambiguity may be too high for many clients. Also, clients and evaluators have to give up control over the process when stakeholders play equally definitive roles at all stages. Clients also have to give up any hope of widely applicable, generalizable interventions to solve social problems, if so much depends on the specific context where the intervention occurs. Many of these ideas are essentially extensions of trends already in motion — such as the movement away from QED; the increasingly active role of evaluation as an ongoing learning process that shapes behaviour; and increasing attention to beneficiary views. The emphasis on negotiation among stakeholders, for example, is not so distant from the "search conferences" approach or, in another context, from the notion of building "shared agendas" among key stakeholders in government R&D labs (Breithaupt 1992).

4. STATE OF THE ART IN PRACTICE

Given the various problems and drawbacks for each method of assessment of the performance of R&D laboratories to which we have referred, the assessment of scientific performance remains more an art than a science. It requires considerable use of expert judgement and qualitative assessment, with various quantitative indicators providing only a partial glimpse into the performance. Because of this, Martin et al. (1987) strongly recommend the use of multiple indicators, each of which provides a partial view of a facet of output. In their view, the combination of the different indicators taken together provides the best approach to judging research performance.

Useful guides to current Canadian perceptions on suitable evaluation models for public research institutions are provided by Breithaupt (1992) and by Neufeld (1992).

Breithaupt (1992) makes the point that evaluators must remember that the laboratory does not control all the instruments to ensure that research results are applied. This is dependent to a large degree on government policies, economic structures, and private sector decisions. However, the lab can attempt to make its research priorities and programs more coherent with national economic policies and the needs of its users. Thus, the strategic plan for the lab should include efforts to secure an appropriate legislative or regulatory environment for its technology mission, to examine the need for industrial standards, to make innovations compatible with existing production systems, and to coordinate its plans with the various industrial development initiatives of the government.

A recent guide to the evaluation of ITRIs by the International Development Research Centre (IDRC 1993) emphasizes the "learning model of evaluation." It lists a large number of suggested evaluation questions. The following list includes only those most relevant:

1. *Institutional motivation:* mission, culture
2. *The external environment:* political, economic, social, cultural, technological level, stakeholders
3. *Performance:* movement towards mission, efficiency, sustainability
4. *Organizational capacity:* program, process and resource management, leadership, linkages.

The research organization's performance is considered to be a function of the interplay of the other three areas: institutional motivation, organizational capacity, and forces in the external environment.

4.1 Institutional Motivation

An evaluation must understand the history, mission/vision goals, and culture that drive an institution's performance from within. An organization's mission statement is a written expression of goals, characteristics, values, and philosophy. Mission statements are relatively new and can be used to communicate to both internal and external stakeholders. There may, however, be a gap between the formal mission and the perceived mission; the stated mission may be dated or misconstrued. One task of evaluation is to assess the congruence of perceived and stated missions. As such statements become common, a new criticism levelled at many statements is that they tend to be banal, are too general, and do not assist planning.

The organization culture is the sum total of values, beliefs, customs, and history. It embodies symbols, myths, and heroes and is often expressed in the collective pride. Organization culture is a powerful force that affects whether members will go all out to push the limits of institutional capacity. It takes time to understand the culture of an institution. Some evaluators use surveys; others use less formal interviews and observations. Irrespective of the method used, the evaluation needs to judge whether the mission and culture of an ITRI are helping or detracting from performance. Some questions that can be asked regarding motivation include:

- How does the organization's mission relate to its goals?
- Are the institution's values compatible with those of its partner institutions and sponsors?
- To what extent have organizational members adopted the mission?
- Is the mission updated and linked to a set of goals?

Some suggested methods for gathering data include asking for existing environmental scans for the institution or related institutions, checking recent studies, interviews and workshops with key informants, and reading contextually (newspapers, magazines, etc.).

4.2 Performance

Measuring performance is conceived of as measuring (i) movement towards the mission (i.e., performance in activities that support the mission — effectiveness); (ii) use of resources (i.e., performance in relation to the resources available — efficiency); and (iii) performance in relation to long-term viability.

Typical indicators of performance include number of publications, number of patents, number of trainees supervised, etc. "Efficiency" is addressed by the amount of external funding received, comparative costs for research, training, and other services, and overhead/program cost ratios. The longer-term indicators are more general and include support earmarked for professional development, quality of working life, institutional innovation and adaptiveness, and institutional reputation amongst key stakeholders.

4.3 Organizational Capacity

Organizational capacity considers a large number of factors under the headings of:

1. *Program management:* planning, implementation, monitoring
2. *Strategic leadership:* governance, structure, strategy, culture, management, niche management, core resource acquisition
3. *Core resource management:* infrastructure, human resources, technological resources, finance
4. *Process management:* planning, problem solving, decision making, communications, monitoring, and evaluation
5. *Interinstitutional linkages:* networks, partnerships, external communications.

As we have already discussed some of the issues in planning and monitoring and structure, culture and management, we briefly review the main points in the other items.

4.3.1 Core Resource Management

Core resource management provides the details of planning and control systems regarding infrastructure, human resources, technological resources, and finance and includes the following:

- Are the buildings and internal services adequate for daily work?
- Is there adequate transportation?

- Are communications systems functional?
- Are there adequate maintenance systems?
- Are the right people in the right jobs? Is adequate human planning occurring? Is equity dealt with?
- Is the institution's level of technology appropriate? Are adequate information technologies in place?
- Is there adequate budgetary planning? Have the finances of previous grants been properly managed?

4.3.2 Process Management

Processes are the internal mechanisms that guide interactions among people for ongoing work. Organizational processes include planning, problem solving, decision making, communications, monitoring, and evaluation:

- Is there adequate (or too much) planning and procedure development?
- Are plans, policies, and procedures generally followed? Why or why not?
- Is the implementation of work smooth-flowing or blocked?
- Are performance gaps and opportunities identified in time to resolve them?
- Are there sufficient members with problem-solving and decision-making skills on the governing board and within senior management?
- Do people feel there is adequate ongoing communication? What are the main vehicles of internal communication?

- If information circulating in the institution becomes distorted, are there corrective remedies?
- Do people have easy access to those with whom they must deal?
- Are there policies and procedures that guide evaluation and monitoring? Are monitoring and evaluation valued as a way to learn? How are data obtained and utilized?

4.3.3 Interinstitutional Linkages

For research institutions, contacts with other institutions are vital. Outreach can be accomplished through networks, partnerships, and external communications:

- To what extent is the institution linked to the external world through collaborative networks? Computer networks?
- Are the networks supported both financially and technologically?

Instead of providing other long lists of issues here, we have provided a few brief statements from a number of different sources in the Annex. The first item in the Annex is a summary of the outline formulated for UNIDO for the evaluation of ITRIs in developing countries. Although this is now dated, it still provides a useful compendium of issues. The second item is the succinct OECD summary of evaluation methods and practices in member countries. The third item illustrates current thinking in Canada, in which evaluation and strategic planning have become embedded together, and the general logic model proposed by the Office of the Comptroller General of Canada for national laboratories. Finally, we provide for illustration and comparison a flow chart of the recommended evaluation procedure in national laboratories in Japan, which has many elements in common with the Canadian model.

4.4 Project Planning and Evaluation in the Private Sector

It may be useful to discuss separately the practices in the private sector. Clearly, business organizations that have profits as their primary motive must feel considerable frustration with the lack of useful measures to evaluate the role of R&D in determining income and profits. Thus, it is useful to examine some of the processes used in large companies to plan and evaluate their research. In the case of one company, Brandenburg and Langenberg (1969) report that project proposals are screened informally before comprehensive evaluation. Promising projects are then subject to formal, detailed appraisal. While a standard format is maintained, the details and precision vary according to the nature of the project and the expenditures. Market factors, economic benefits, and the criteria for technological success receive special attention. The analysis of market factors imposes a discipline on the scientist, the division, and the R&D manager to consider what purpose the product will serve and who will buy it. While the criteria for inclusion are based primarily on return on investment, they include judgemental factors as well, such as how the project relates to other projects and how it could strengthen the technological success of another project. "Selection decisions have not been reduced to a precise formula."

Allen (1970) surveyed 112 industrial organizations. He found that a commonly used system shared by many organizations begins with the filling out of a proposal form. The form is submitted to a committee that may allow the project to proceed up to a certain estimated cost. Larger-scale projects are referred to the executive board for financial approval subject to their congruency with the long-term interests of the organization. Evaluation is carried out at a high level, with a research member on the executive board.

In Allen's study, 60% of respondents had made provision to allow researchers an amount of "spare time" to work on their own ideas. Only 11 put the use of this spare time totally at the discretion of the individual researcher. However, in 62 others, the R&D director or project leader influenced the decision regarding the use of spare time, and 38 organizations had specified time or monetary limits on a project, beyond which level the project must be submitted to the evaluation procedure. Projects commenced in the researcher's spare time, if approved, became a part of the ongoing research program. In other cases, no formal resource limits were specified, but levels were set at the discretion of the research department head. In general, the R&D director/project leader handled the administration of individual research, and projects were informally assessed.

The information collected for evaluation was separated into the procedures for approval and the information necessary for implementation. Only 16 organizations used none of the methods of net present value, payback period, or average rate of return on capital invested when assessing commercial viability of projects. Of these, 14 went on to estimate "probability of commercial success." Some estimates of the probability of commercial success, technological feasibility, and probable development costs were made by 99 organizations. One-third of the respondents had a single proposal and evaluation system that was used to evaluate all projects, while others had different systems for basic/applied research versus development projects. Only 12 applied any numerical weights to their criteria.

About two-thirds of respondents reviewed basic research projects at least twice a year, and several reviewed them every month. Important review criteria were any changes in the chances of technological success and of commercial potential and new estimates of further development and completion costs. In detailed discussions, many stated that because an R&D department carries out a variety of long- and short-term activities, such as "trouble shooting," quality control, and providing technological services, a single objective is difficult to define. Many also stated that the poor definition of project objective often leads to difficulties in carrying out the research. However, they still felt that although criteria are difficult to establish, they are necessary to assess progress. In many organizations where formal methods were used, a problem was that not enough new ideas were being submitted to provide the variety needed to select an appropriate portfolio.

Virtually all organizations cited problems in the coordination and communication between R&D and other departments. While the levels and structures of interaction varied widely and interactions appeared to be a function of the individuals concerned, as well as the management and organizational structure. Cases were found where R&D and production were located in the same building to increase communication, but at the same time a lack of formal channels hindered communication. Most R&D staff interviewed expressed great interest in the overall activities of their organization and wished to see successful products stemming from their research. Dissatisfaction was noted within certain organizations that did not inform researchers of the value of end results. In many cases, market conditions were not adequately assessed to enable potential benefits to be estimated, and market changes were not always noted, resulting in lost opportunity. Scientific personnel seemed to prefer well-defined projects with fairly precise objectives, whose commercial potential was known to them.

Agarwal (1974) reports that at Kennecott Corporation the process starts with an understanding of the company's business, its market, and its technological capacity to develop a mission-oriented research program. Large projects with good probabilities of success are kept as a major part of the program. Identification of the specific research tasks is followed with a schedule and an appraisal of skill availability to ensure success. The large projects are evaluated by engineering, operation, finance, and marketing as early as possible. To minimize business risks in R&D, there are annual meetings with division presidents in order to gain approval for project objectives; meetings are continued with divisional people to monitor progress and adjust direction when necessary. While projects are planned several years in advance, they are reemphasized or eliminated according to whether technological and market assumptions hold true over their life spans. Significant milestones are chosen, and decision points and criteria are decided. The technological and supporting manpower requirements are forecast to allocate manpower and avoid peaks and valleys. For process development projects, their impact is estimated by asking how much sales the new processes affect. What are the current costs, expected costs, and resulting annual savings and their significance? For instance, in new processes, Kennecott requires that the minimum savings should be $1 million after taxes to qualify as a large project. No discounting is taken into consideration initially. Useful new products are screened keeping in mind the 15% minimum return on investment.

The lack of use of formal models in industry, although many such models are available, is confirmed again by Fung and Shapiro (1976). They also report a 1973 study that found that 32% or 64 of 200 Canadian firms did not use *any* project selection techniques. Ranking and economic methods, but not mathematical ones, were popular among users. They also report the view that many found research planning becoming a very resource-intensive process. In one company, it took 10 years to formulate a corporate research plan. However, they do not provide information on whether the resultant plan and the research were then evaluated and if the resultant benefits justified the resource cost of planning.

Cutler (1979) describes a matrix process used by Whirlpool company and cautions that evaluations of the economic value of the final outputs must be undertaken with sensitivity to the possible changes in many factors beyond the researchers' and the company's control. He also describes the role of technology forecasting groups, which serve to bring wider technological developments and long-range influences to bear on the decisions. Danila (1985) reports on the evaluation procedures used by a French pharmaceutical company and highlights the role of multiple committees and criteria in use. She finds an important role for return on investments as a criterion but no use of formal models. Some other reports of evaluations in industrial settings include Clogston (1982) on the use of evaluations at Bell Labs and Porter (1978) on their use at Mobil. Mobil Research has undertaken post-audits of applied R&D since 1960 using the same technique to predict the value of proposed projects. They find that post-audits motivate researchers to select and complete money-making projects, help identify high business risk projects, demonstrate research productivity, identify productive research areas, and increase confidence in predictive evaluations as a tool for guiding research.

Quantifying benefits emphasizes the profits rather than technological achievements. Publishing the numbers motivates researchers who have been successful and encourages them to complete projects that may have lost their technological challenge. Quantification of benefits at Mobil revealed that many projects failed owing to excessive business risks associated with (i) capital outlay, (ii) being outside the mainstream of the company's business interests, (iii) poorly defined targets, and (iv) small-volume products and processes not attractive to the market. Additionally, corporate management is more aware of the uncertainties, as demonstrated by the wide fluctuations of benefits on a year-to-year basis.

While rates of returns are commonly estimated in many private ITRIs, Collier (1977) warns that a traditional but superficial question asked of R&D is to measure the value of corporate research operations in terms of how much money has been made from R&D. He believes that this asks about the overall system performance of the company, of which R&D is just one part. Profits will of course depend on the appropriateness of research objectives made by management. However, they depend as well on how well they are interpreted, the quality of engineering designs using R&D's technology, the effectiveness of how the idea is translated to manufacturing, and how salespeople promote the product or service. With lags of up to 15 years from the initiation of research to commercialization, feedback on economic impact is too slow to be relevant. Other traditional measures are based on counting patents, papers published, or engineering drawings. These indicators may not reflect the goals of the company. A system of performance measurement needs to look at R&D independently from the rest of the company and to do so quickly enough to provide useful feedback. Given the many difficulties of measurement, Galloway (1971) makes a virtue of keeping the evaluation criteria simple. Liberatore and Titus (1983) report on the techniques used by 29 firms and find that one or more financial evaluation methods such as Cost Benefit Analysis, Net Present Value, and payback calculations dominate in firms and are used by almost three-quarters of the firms.

In a recent review, Cordero (1990) notes that both quantitative and qualitative measures are imperfect. Some of the quantitative measures miss important facets of outputs and resources, but to correct this with qualitative measures may be expensive. A study by Booz-Allen (1982) found that 65% of over 700 manufacturing firms used formal measures to evaluate new product performance. Schainblatt (1981) found seven of 34 innovative firms using quantitative measures of resources and outputs routinely. Another four used quantitative measures of output only, and one of the two occasionally. Another study of 124 R&D managers from 40 industries asked about the use of 13 facets of output and showed that the most frequently employed criteria were quality of technological output, unit degree of goal attainment, and amount of work done on time.

4.5 Public ITRIs and the Private Metric

Burgess (1966) states that the yardsticks employed in evaluating R&D programs in industrial firms cannot be used for government laboratories. This paper discusses the problems and points to factors that must be considered in the government context. Government laboratories differ from both industry and academic labs, as government is called upon to assist administrators and policy makers in regulatory work that one rarely associates with the concept of a research laboratory. Araoz (1993) also cautions that among the valid contentions of public sector ITRIs, one is that many of the activities that they are mandated to undertake are not amenable to measurement by a market metric. They are often required to provide advice to governments on certain technology issues, to develop standards and specifications, to increase the trained research manpower, to keep abreast of relevant international technological developments, and a number of other needed activities that do not often have revenue potential.

From his interviews with research managers of ITRIs, Araoz (1993) states that effectiveness remains difficult to measure even though it is notionally simple. Even in a very concrete case where an ITRI has developed a new product and a manufacturing plant is opened to produce it, it is difficult to allocate the resultant economic benefits between the research institute and the other investments of resources. Any formula is necessarily arbitrary.

Araoz reports on the frustration of many public ITRIs in developing countries with regard to implementation. In most cases, the commercialization stages, after the initial product development, are the responsibility of other public and private sector organizations. The Caribbean Research Institute, in Jamaica, felt that in some cases the other government organizations did not approach commercialization correctly. There is often inadequate promotion, inefficient production, and insufficient attention to market demands. CARIRI also cite similar marketing failure even with private sector organizations that have taken up licences. In a number of cases in India and Uruguay reviewed by the author, similar problems have been cited by ITRI managers. Because of such instances, many ITRIs have additionally taken on the product development to commercialization tasks themselves, often with new subsidiaries whose mandate is for the specific commercial activities.

4.6 Concluding Remarks

To summarize, institutional evaluations typically gather information through observation, interviews, and perhaps small-group discussions. To improve reliability, contrasting sources of information are sought — interviews, for example, should tap not only internal sources but also mission personnel, beneficiaries, staff at other institutions, and so on. To reduce the risk that external evaluators are perceived as threats, some institutions specify procedures that make the evaluation a joint effort involving internal management, such as the use of steering committees.

Nonquantified approaches are dominant, and, even where numbers can be counted (e.g., number of projects completed, number of staff trained, number of patents, number of publications), their interpretation is still subjective. Case studies are relatively common and should include some thoughts on "lessons learned."

A common framework in the evaluation of projects and programs distinguishes among (i) rationale (the fit between the program/project and larger goals); (ii) effectiveness (achievement of objectives); (iii) efficiency (optimal use of resources in achievement of objectives); and (iv) impacts and effects. Institutional evaluations should include discussions of additional factors such as the institution's external environment, organizational approaches such as performance measurement, financial management, program management, decision making, communications, and linkages. However, these elements are not always included because of potential fuzziness in the elements. This fuzziness aggravates a tendency to attribute a problem to poor internal management when in fact its origins are not under internal control. Because external factors are so important for public ITRIs, any evaluation approach that focuses only within these boundaries will ultimately be inadequate.

Fifteen papers on state-of-the-art R&D evaluation in Europe are summarized by Shankar (1985). Methods used vary between and also within countries. The consensus is that evaluation methods, techniques, and criteria need to remain flexible to adapt to particular programs. It is also stated that timing is important: an early evaluation could lead to erroneous results on impact, while a late one may not provide information for interim decisions.

A special feature of Italy's R&D evaluation is that it is integrated into management in a more or less permanent way. Denmark's evaluation is an integral part of the research system, even in hiring scientists, promotions, publications, grants, etc., and evaluators come from a variety of disciplines. The Netherland paper states that evaluation is more of an art than a skill and that implementing evaluation findings can be more difficult than the evaluation itself. Additionally, bibliometric evaluation is found to be reliable for basic research, although complementary peer advice is desirable. In applied research, these methods are inadequate. Feedback from evaluation users is necessary to improve the evaluation methodology. The evaluation team must be independent of the program management team and composed of scientists, economists, and potential users of R&D results.

For the purposes of the Latin American comparative study, two additional points bear repetition. First, the propensity to produce papers, reports, or patents varies by institutional objectives and mandate, by scientific field and area of application, and by sociocultural factors. Second, when we are concerned with developing country institutions, these indicators can be plagued by systemic underrepresentation and simple nonavailability. It is worth mentioning here that the problems of using some of the quantitative indicators of S&T outputs in developing countries are well known and often discussed. However, a recommendation that can well be pursued by this group is that instead of continuing to wring their hands about this state of affairs, a constructive option would be for Latin America to build up the data bases by which regional research performance can be better assessed.

5. MANAGING THE ITRI

If we could measure the inputs and outputs of ITRIs unambiguously, we could leave the internal issues of management styles, procedures, and structures alone to those whose job it is to manage. Unfortunately, however, as we have said, it is not possible to do so in a relatively unambiguous way. We also know from sociological and managerial studies that how the black box of the R&D system is structured, organized, and managed plays a critical role in transforming the inputs into valued outputs.

Ritchie (1970) feels that there is an overemphasis in the literature on the *evaluation* of innovative ideas rather than the *generation* of good ones. Landenburg (1969) notes the need for R&D managers to be receptive to new ideas and to pursue them. Rockett (1970) recommends that a special group be responsible for evaluations to avoid the termination of ideas at early stages by unreceptive managers. At a large U.S. corporation, Baker et al. (1967) found that ideas were associated with the combined recognition of an organizational need, problem, or opportunity that is perceived to be relevant to the idea, and of a means or technique by which to satisfy the need, solve the problem, or capitalize on the opportunity. The authors noted that three-quarters of the ideas presented were prompted by a knowledge of an organizational need, and only one-quarter by the knowledge of capability. Of the 271 ideas studied, 47 were judged by the "idea generation group" and the lab director as "best." Of these, 40 were the result of perceiving a need. Marquis (1969) found that about three-quarters of the 567 improved product/process innovations he surveyed were stimulated by market demand or a production need. The recognition of demand is a more frequent factor in successful innovations. Achilladelis et al. (1971) found that successful innovators pay much more attention to marketing and have a much better understanding of user needs.

Although the stimulus for an idea appears to come from outside, the source of ideas is primarily internal. In a study of 34 small electronics manufacturing companies, Mansfield (1971) found that 89% of the labs' work was a result of internal ideas identified primarily (62%) by the R&D staff. Management is more likely to recognize a new idea and to reward its originator if it is considered relevant — i.e., if it satisfies an existing need, solves an existing problem, or can be developed into a new project that is compatible with the organization's overall goals and objectives — and can be investigated with existing resources and facilities.

Baker et al. (1971) found that ideas that rated highly with researchers were more predictable and had shorter time horizons. Management rated urgent projects more favourably than researchers. Avery (1959), on the other hand, found that in seven of 10 R&D labs, what is the "best" idea changes at different levels of the hierarchy. Martin (1967) found that ideas will be rated higher:

- the lower the perceived probability of failure,
- the lower the perceived cost of implementation, and
- the more urgent the problem.

There are plenty of papers and books on how R&D should be organized and carried out, but there is some evidence that these techniques are not very frequently applied in practice (Freeman 1969, p. 26). However, some findings appear to be relatively robust and are guiding the current phase of reorganization of ITRIs. To the extent that the current forces for reorganizing the ITRIs seek greater economic benefits from innovation, a recent study of the best practices regarding industrial research carried out for and in collaboration with companies and ITRIs, based on evaluation of 84 successful European projects, is relevant. It finds three successful types of project-initiating practices:

1. good practices on stimulating demand for new projects,
2. transferring company requests into research projects that may involve several companies in execution, and
3. practices on evaluating project ideas.

The study found that the practices on project operation reflect the need for active company participation in establishing priorities. The goals should be defined in terms of clear benefit to the company rather than in general research terms, with clear economic focus to operate on the sales activity. Project implementation practices are divided into information activities (demonstration projects, information events, publications, software programs, training courses) and accompanying actions for bringing the techniques into use (company-specific training, transfer of people, counselling by ITRI).

General conclusions are that successful research projects are not merely products of science pushed by Research or demanded by a request of client companies. They are rather *products of circular processes where continuous dialogue occurs* between the two parties. The operational aspect of collaboration is primarily a social process and secondarily a technical process. Cooperative research needs to be closely interrelated with practical implementation of the findings at the "shop floor level," with focus on implementation. Practical implementation often includes aspects of human resources, such as changed organization of labour and upgrading of skills. Such research may be aimed at immediate industrial use or can be of a preliminary nature with indirect impact on companies. Of the research projects aimed at immediate implementation, there are those resulting in profits or cost reductions for companies, those aimed at welfare and other nonprofit issues such as the environment or safety, and those aimed at indirect aspects such as improved intercompany cooperation and training. For the majority of cases studied (up to 80%), ready-to-implement outcomes were the objective. The others were aimed at "upstream research," where the issues needed further elaboration, or at building interest amongst a group of client companies.

The types of enterprises involved in the ITRI client group include technology-driven companies whose goal is to be in the forefront in identifying and bringing into use new technologies, active followers, passive followers, and technologically inert companies. A schematic model of corporate research is presented by FEICRO (1993), which begins with idea generation, followed in sequence by project design, project execution, knowledge transfer, and finally industrial implementation. We must recall here that the process is in fact highly circular. The model for cooperative research is viewed as a communication and learning process with both direct and circular interactions between the ITRI and the client.

The challenge of successful implementation depends almost exclusively on the user having learned the necessary skills to operate the new technology, a process that must be integrated into the research process, so that the user already knows what to do when the results become operational. It is also dependent on active involvement of the client. Management performance focused on improving the use of technology is thus crucial. The FEICRO study found the product/market mix of the successful ITRIs to be made up of:

1. Products that serve as a direct link to firms such as consultancy, testing, and trouble shooting, which are common problems in a particular branch of industry. These often lead into cooperative longer-term research projects.
2. Applied research projects oriented to improve production processes and/or serving as a basis for new/improved product development.
3. Strategic research aimed at continuous refreshment of knowledge and incorporation of new technologies/knowledge in the ITRI.

Good project management and control systems within the ITRIs are necessary in order to follow the progress of the project with respect to the time frame agreed upon with the clients and the financial progress made. The cooperation between a sector of an industry and an ITRI is predominantly one of partnership, especially when contracts are based on membership organizations. However, ITRIs must remain aware that the borders between many sectors of industry are fading. A number of cases discussed concern transfer of technologies. For this, the value of a full-scale demonstration facility is emphasized. After an adoption decision by a firm, the ITRI that could provide continued technical and counselling assistance was more successful.

In analyzing the good practices on stimulating demand, FEICRO (1993) found that idea generation often came as a spinoff of current testing and certification activities at the research centre. Testing was found to be the most important instrument for the continuous interaction with client group companies.

A Spanish research association discovered problems with a product, as most companies' products had been tested at the centre. The centre was able to take immediate action by inviting company representatives to participate in a project for determining the solution. A similar situation involved wood-based windows in Denmark. A Danish product certification scheme formed the basis for identifying the problems of the producers, and the ITRI went on to develop and transfer the solution.

Demonstration projects are especially effective in convincing target group companies of the possibilities of a new technology and often provide a good opportunity for involving companies actively in the full-scale development. Collaboration with client companies should not be confined to "stand alone" projects but rather to a successive formulation of interrelated projects — one project generates the next.

One good practice on appraising project ideas is to evaluate the project idea with client group representatives, and this practice is employed by virtually all research centres studied. A clear definition of the objectives of the project remains important in getting approval of the client base. Here, economic gains/reduction of costs should be elaborated. The initial plan guides the general development of the project, but it should not be inflexible, as continual adjustment is necessary to cope with emerging issues. High involvement of the ITRI in the early stages of the project is essential. Joint project operation is also recommended. Good project management practice includes reporting of the results, including financial data, to the clients on a regular basis, often every three months. An instrument for enlarging the group of companies is to establish reference groups or informal discussion groups that can be used as fora for testing of preliminary project results.

Intermediaries linking ITRIs and user companies can be consultants and research foundations.

Some of the knowledge transfer means that were found useful are:

- "hands on" demonstrations (in the laboratory or full-scale exhibitions),
- publications (reports, leaflets, handbooks, computer programs),
- meetings (regional or project bound),
- training (workshops, seminars, knowledge module in courses),
- company-specific actions (company visits, counselling, transfer of people) for when implementation demands radical changes in the organization of work and in skills requirements,
- rules of operation for parties (technical norms, specifications, certification schemes).

Given the very large number of factors that the literature cites as important, we will not be able to discuss them further here. Other useful sources include Clarke and Reavley (1988), who discuss many factors that have been considered relevant in Canadian labs, and Jain and Triandis (1990), who provide a good discussion of what are considered good management practices.

6. SOME OVERALL TRENDS IN PUBLIC RESEARCH INSTITUTES IN OECD COUNTRIES

. We have noted already that there has been a considerable rise in interest in appraisals of the roles, missions, structures, and organization of public research institutions in all OECD countries. This rise in interest has been motivated by a combination of budgetary pressures, ideological shifts, and rapid changes in the directions and pace of S&T and its application in economic activity, as well as a desire for greater effectiveness and in particular for greater economic utilization of available expertise and knowhow. OECD (1987b, pp. 29-31) lists major exercises undertaken in the United Kingdom, Denmark, Canada, Norway, Germany, Australia, and the United States between 1979 and 1986. It points out that one can have too many appraisals, as in the United States, where, over five years, there were nine major reports that covered the federal laboratories under the Department of Energy.

All the reorganizations that have been implemented have the following main characteristics in common, although there are important variations in the details of implementation and design in different countries:

- institutional autonomy,
- decreased public and/or institutional funding and increased contract and program funding,
- greater user interaction and feedback,
- greater attention to utilization.

All of the above are interrelated to a significant extent, but for clarity it will be useful to deal with each one separately.

6.1 Autonomy

The need for autonomy of research and researchers and the incompatibility of the nature of R&D activities with tight administrative and financial systems and public sector rules and procedures are well recognized in the sociological and management literature on scientists and scientific institutions. Research institutions often find the rules of the public sector too rigid in terms of accounting and staff procedures, while at the same time the goals and objectives of the political system shift too quickly to permit effective longer-term research plans.

The need for autonomy is increased further if the institution is to interact in various ways with the private sector, market more of its knowledge and services, and pay greater attention to utilization. RockCliffe Research (1992) discusses a number of models for greater autonomy of public research institutions. These include The Government Owned Contractor Operated (GOCO) model in the United States, which is an older model and predates the current trend. The Executive Agency (EA) model being developed in the United Kingdom will remove civil service restrictions, and the EA can market its services to both the government and the private sector. The ownership of the EA remains with the government.

The same report discusses at length a number of Canadian models, which are also reviewed by Mullin (1993). It discusses handing over industrial research institutes to appropriate industrial associations with the costs shared between the government and the industrial association as one model. The two Canadian examples — PAPRICARN and FORINTEK — are in the pulp and paper industry. Although the report suggests that many believe this model has not improved the effectiveness of FORINTEK, that may be due to the underspending by Canadian industry on in-house research.

RockCliffe Research (1992) reviews the model of a public sector corporation and suggests that such a move would be better for the National Research Council of Canada. However, it points out that even then the corporation would be hobbled by many rules limiting its ability to earn revenues from successful technologies. It has a number of other models of privatization, which range from outright sale, to employee takeovers, to contracting out of technological functions, variants of the U.S. GOCO models, and, finally, to the leasing of equipment and facilities. Most of these models have been tried for some of the federal and provincial research institutions, but we do not have enough information to comment on the outcomes of these changes.

6.2 Funding

Without a change in organization and/or laws, it is often necessary for government-owned institutions to deposit all self-generated revenues back into the general revenue account. This certainly provides no incentive to the scientists or the institution to undertake the difficult and scientifically unrewarding task of technology transfer and commercialization. "No model will promote successful commercialization unless...the labs are able to benefit from the commercial success" (RockCliffe Research 1992, p. 72). A Canadian proposal that 20% of self-generated revenues be retained by the research institutions is criticized by

RockCliffe Research as inadequate. RockCliffe Research recommends that research institutions be allowed to retain 100% of net revenues.

In this context, the German policy and experience with the Fraunhofer Gessellschaft (FhG) is noteworthy. The FhG is a group of 34 state-supported industrial research institutes. In 1973, there was a review of policies towards state support for the FhG. A change was instituted whereby the amount of state financial support was matched to the amount of contract research that was generated by the FhG. Thus, the more money the FhG received for contract research, the greater the state support. In essence, while Canada debated a policy for 20% retained earnings, the FhG had operated for a decade and a half on the basis of 200% retained earnings!

Schmank, in his review of the FhG experience in Germany, points out that in 1975 the FhG had a budget of 112 million DM. This grew to 511 million DM in 1986, and during the same period its staff increased from 456 to 2680 (Schimank, p. 223), all due to the growth generated from contract research. In the end, the FhG was so successful at attracting contract research revenues that the state component of its budget could not keep up its share owing to restrictions on public finance. In recent years, the state share has dropped to 30% of the total budget for the FhG. This performance would definitely be considered a great success. However, Schimank cautions that this new policy was *not* in fact due "to a carefully designed (government) technology policy" but a direction that FhG had itself earlier initiated and moved towards (p. 223). There is also another caution that with the new overwhelming predominance of contract research, perhaps more medium-term directions for innovations are being lost sight of.

In the Netherlands in 1985, the TNO as a whole was receiving almost 60% of its total funds from contract research. Within the TNO research institutions, the percentage varies widely, ranging from a high of 95% for some of the ITRIs to a low of 10% for institutions whose primary mission is in social sectors and whose primary users are in the public sectors (OECD 1987b, p. 37). It should be emphasized here that not only will these opportunities vary by sector of operation, type of ITRI, and types of research activities, they will also vary from country to country and within the country by regions, based on the size, strength, type, and structure of industrial production. So, while the ratios are not automatically translatable across countries and especially to developing countries, the general principles should remain valid.

For instance, RockCliffe Research, in its review of models and policy for publicly funded R&D, points out that for Canada an important feature is the lack of export-oriented, indigenous, technology-intensive firms. Thus, Canada has a weak demand and receptor capacity for technology from national ITRIs. RockCliffe Research goes on to state that because of low levels of industrial R&D at the firm level, Canada has a high percentage of its scientific manpower in the public ITRIs. While this makes it urgent to improve the outputs and effectiveness of this sector, at the same time it is more difficult to achieve economic results by working on the supply side alone. These observations would probably be even more acutely valid if we replace Canada with any of our case study countries.

6.3 User Participation

The principle of the importance of user involvement in the work of the R&D scientists and institutions has been well recognized for many decades and has been discussed at several points so far. In spite of that, many reviews of ITRIs in developing countries and also of research institutions in industrial countries, bemoan the very weak links between the users of the research and the producers. In many cases of earlier attempts to involve users in the directions, plans, and choice of research, the emphasis had been to place industry (for ITRI) representatives in the top governing boards of research institutions. As utilization and dissemination of research gained prominence in ITRI objectives, industrial extension, outreach programs, conferences, newsletters, and so on began to be developed. In this regard, the almost three-decade-old Industrial Research Assistance Program (IRAP) of the National Research Council of Canada is an excellent model. It has been assessed recently and found to provide a good return on resources used. There have been many different initiatives in Canada and other OECD countries to involve users in new and different ways, at various levels, and at different stages of the research process.

6.4 New National Directions for ITRIs

6.4.1 Canada

Mullin (1993) examines some of the different models of institutional organization and funding mechanisms that have been attempted to reduce the sources of inefficiencies in public ITRIs in recent years in Canada. He says that in recent years the Canadian policy towards technology development and for the ITRI has been motivated by the desire to improve international competitiveness.

The new policies aim to ensure that public laboratories, whose primary task is towards industrial support, cooperate increasingly with their clients, emphasize technology transfer activities, and examine the scope for privatization of public facilities. An important inducement provided in 1986 for Canadian laboratories to become commercially oriented has been the provision that in addition to their government budgets, they can, for the first time, keep a share of all revenues directly generated by them. Until 1986, any revenues earned by the federal laboratories went to the general revenue of the government, providing weak incentive towards commercialization.

6.4.2 United States

For the United States, Crow and Bozeman (1987) note that widespread changes in organization forms and environments. They suggest the emergence of new forms of organization have blurred the distinctions between the government-private organization and cooperative and interorganizational research has proliferated. Yet, unfortunately, public policies are often made on the basis of outdated assumptions about competencies, liabilities, and performance of labs. Crow and Bozeman (1987) also caution that strong market influences reduce stability and program horizons and induce frequent changes in goals. So, for labs directed to produce generic knowhow broadly for industry, a redirection to the market may be ill advised. What is required is the requisite variety among the group of R&D institutions rather than that they should all be of one type. However, Crow and Bozeman (1987) also find that a reduced level of public support for labs producing generic knowhow can force greater links to the market, shorter planning horizons, and more applied research products.

6.5 Australia: The Review and Reorganization of CSIRO

In 1985, the Australian Science and Technology Council (ASTEC) was requested by the Prime Minister to review Australia's Commonwealth Scientific and Industrial Research Organization (CSIRO), which had already been extensively reviewed previously (Government of Australia 1977). Oldham (1988) describes the process and discusses some of the main issues that arose in the review of CSIRO. The 1977 review of CSIRO had suggested that the divisions of CSIRO be reorganized into semiautonomous institutes, partly discipline and partly sector based, to be managed by an Executive Council, including members internal and external to CSIRO.

The 1977 review made 120 recommendations, of which 108 had been accepted by the government. So, in 1983, the government undertook another review to determine the implementation of the recommendations. This review found that the institute system was not working well, the institute directors had not been made members of the Executive Council, and the new Advisory Council of 31 members had such a diverse background that they could hardly agree on anything (Oldham 1988, p. 2) -- and when they did agree, their advice was usually ignored.

CSIRO felt that it had made the necessary changes, but both industry and government felt that the changes had been insufficient. While some changes had been made in parts of the organization towards a business orientation, they were marginal. Those research groups that were able to obtain money for research from industry were frustrated, as they could not retain the additional funds. In the interim, Australia's economic position had worsened considerably. The Department of Finance was concerned at the substantial outflows of resources to CSIRO for what they felt were paltry returns.

Oldham (1988) describes the procedures followed in the ASTEC review of CSIRO. Essentially, a committee was formed with members from within and outside the organization, having a mixture of academics, research directors, and industrial users. The methods for collecting information included submissions from interested parties, of which over 300 were received; visits to labs and discussions with chiefs and staff; interviews with research managers in other countries; and literature reviews. Directions of work, tentative findings, and conclusions were reported at open monthly meetings to ASTEC. This allowed for continuous feedback into the review process, which was completed over a seven-month period (for more details, see Oldham 1988).

Some of the key elements used in the review included the involvement of diverse groups of stakeholders and the organization and a commitment to keep the process open and transparent. The review committee wanted to share its results with CSIRO throughout the process in order to increase the likelihood of adoption of the final recommendations.

The key recommendations that emerged from the process included the following (p. 13):

- researchers should seek more contract work from industry, be more applications oriented, and ensure effective technology transfer to end users,
- a Board should be created for overall management,

- earnings should be retained while government support remains unchanged,
- the link to public service should be eliminated,
- goals should be developed in terms that can be evaluated at predetermined points,
- public and private staff should be exchanged to improve communications,
- appraisal methods should be used to recognize achievements, and life appointments should be reduced,
- facilities should be made available for training,
- there should be technology training for commercialization,
- opportunities to establish independent and joint venture companies should be sought as a way to commercialize CSIRO's technology.

The above recommendations can be summarized to consist of dimensions creating greater autonomy, reduced role and demands from the public sector, greater links to and guidance from the market end users, increased emphasis on communications and dissemination, and a new, applications-oriented ethos.

Evaluation Outcome

CSIRO responded with an internally generated proposal for restructuring. The new CSIRO Board rejected it and appointed a consulting company to advise on the structures. Five business sectors were identified to cluster activities, and one new one was added for emerging information technologies. The original 41 divisions were allocated among the six new institutes. A business system was developed to guide the institutes' operations, and each of them is now required to obtain 30% of its budget externally. Priorities are defined, with new advisory committees operating at division or institute levels, and each institute has its own planning officer. If reorganization continues as planned, 450 of 700 staff will be eliminated.

These recommendations were not without criticisms, as we can anticipate from the earlier review of the literature. Some critics of the recommendations felt that they overemphasized the role of short-term, contract research for industry; that they did not provide a strategic plan for CSIRO; and that they had too little reflection on the future role of government policies. Some pointed out the problem of planning for research over a 10-year horizon when the government changed in three years.

Other problematic questions remain, such as reconciling institute priorities with national research strategies. Measuring a scientist's or engineer's contribution to industry may become more difficult because of the confidentiality of much information. With greater involvement with commercial contracts, the problem of intellectual rights will increase. Links between government departments may become strained, and scientists become frustrated by secrecy imposed on information owing to commercial considerations. The present institute structure will likely impede interdisciplinary projects. In fact, one Board member felt that the new structure "locked the organization into feudal baronies."

A 1987 survey of the professional staff found that 95% accepted the changes, accompanied by resentment of the lack of communication concerning them. It was estimated that 10% of the staff were very opposed and felt bitter or betrayed. Many felt that it will be very important to ensure that innovativeness not be quenched by a ruthless application of the business systems approach. One director spoke of the contrasts in direction with an example of research in superconductivity that enabled materials to become superconductors at liquid nitrogen temperatures: a scientific ethos would lead to searches for new materials that become superconductors at even higher temperatures, and a business direction would emphasize applications of currently known superconductors at less convenient temperatures.

It was recommended that the level of appropriation funding for CSIRO be maintained, but the recommendation was not accepted by the government, and cuts were received.

Another recommendation suggested that the recognition of achievement not be dependent solely on publication. This was welcomed but proved difficult to put into practice.

Oldham (1988) concludes that the purpose of the ASTEC recommendations was to aid wealth generation, increase technological infrastructure and capability, and better manage known problems. The review itself endangered a great amount of scientific research, lost during the time taken for the review and government response. However, this real loss must be compared with the potentially greater loss that may have resulted without the reorientation. Without the review and subsequent reforms, CSIRO's budget would surely have been reduced. It was also reduced after the process, but the cuts could have been greater.

The review was carried out openly and with room for input from everyone, but still many research staff felt uninvolved. Along the way, key officials were told how the conclusions and recommendations were developing. This process itself had its own impact on policy, irrespective of the final report. The greatest impact of the study came from a few important recommendations, suggesting that many evaluations can make their reports shorter and more coherent, emphasize the key points, and provide fewer recommendations.

CONCLUDING REMARKS

We have noted the ironical situation whereby the assessment of the quality and quantity of S&T remains more an art than a science. We have discussed and highlighted that, even though there is no unique recipe for assessment of ITRIs and that each assessment must be tailored to the specific purposes, nature, and mission of the ITRI, as well as its wider industrial, economic, and technological environment, certain general principles and approaches are now commonly accepted.

Important common procedures include the use of review panels with both subject or area expertise and evaluation expertise, the importance of listing and documenting different types of outputs, the uses of different quality and value indicators, and, finally, the critical importance of user interactions of varied sorts. In current idealized practice, there is felt to be a need for close systematic and overlapping interactions between strategic planning for the ITRI; initial target setting; the operational procedures for choice of activity, methods of work, project selection approaches, and ongoing monitoring; and periodic ex-post assessment of work done and its value and impact. Finally, given the various problems of assessment of R&D outputs and its importance, we believe, with Martin et al. (1987), that carefully matched institutional comparisons can be an extremely valuable tool and should be attempted more often. It will be very interesting to observe the extent to which the current set of case studies allows for such comparisons.

We have also suggested that, ultimately, it is not the evaluations of ITRIs that are creating the pressures for change, but that the evaluations are the result of new pressures on ITRIs. A new consensus, even among governments not wedded to the market and inclined to intervention, is one highly critical of many existing public organizations and services. It is prepared to reduce government support where it is deemed to be ineffective and open up such areas for the private sector (Yencken 1987). Another consensus among science policy analysts is that the high growth phase for scientific research in the 1960s and 1970s has levelled off to a "steady state phase" (OECD 1987a; Ziman 1987). Schedvin (1982-83) discusses the cycles of government science between applications-oriented periods during times of economic decline or war and science-oriented periods during times of economic prosperity. This requires greater selectivity and concentration of the scientific enterprise than earlier, with new organizational structures and management modes and more rigorous evaluation of outputs (Oldham 1988, p. 5). Further, in many of the peripheral countries, a new pressure group that has emerged, with the low growth or declining resources for S&T and the increase in total national research capacity, consists of university researchers (Oldham 1987, p. 5). In many of these countries, the capacity of university research centres has grown in the past three decades, and they are reluctant to automatically cede resources to the national ITRIs. The final driving force for change is generated by the rapid changes in technology, which have made many traditional practices obsolescent.

So, while there are many microlevel recommendations for improving the management of ITRIs and thereby their performance, most of the new recommendations focus on changing the environmental and organizational context of the ITRIs. In recent years, the major policy instruments used to reform ITRIs in OECD countries have included greater autonomy, reduction in public sector norms and regulations, increased reliance on generating their own funds, greater involvement of users, and a greater emphasis on utilization. For each one of the above, a number of methods have been found to be useful and may be relevant to the Latin American countries.

REFERENCES

Abend, C.J. (1985) "The Corporate Black Hole," **Journal of Product Innovation Management**, Vol. 2, No. 2, pp. 88-92.

Agarwal, J.C. (1974), "Project Planning at Kennecott," **Research Management**, Vol. 17, No. 3, pp. 34-37.

Ali, A., Kalwani, M.U. and Lovelock, D., (1993), "Selecting Product Development Projects: Pioneering Versus Incremental Innovation Strategies," **Management Science**, Vol. 39, No. 3, pp. 255-274.

Allen, J.M. (1970), "A Survey into the R&D Evaluation and Control Procedures Currently Used in Industry," **Journal of Industrial Economics**, Vol. 18, No. 2, pp. 161-18.

Allen, J.M. and Norris, K.P. (1970), "Project Estimates and Outcomes in Electricity Generation Research," **Journal of Management Studies**, Vol. 7, No. 3, pp. 271-287.

Allen, D.H. and November, P.J. (1969), "A Practical Study of the Accuracy of Forecasts in Novel Projects," **Chemical Engineer**, No. 229, pp. 252-262.

Angle, A.H.L. (1989), "Psychology and Organizational Innovation," in **Research on the Management of Innovation: The Minnesota Studies**, A. Van de Ven, A.H.L. Angle, and M. S. Poole (eds.), Harper Row Publications, New York, Ch. 5, pp. 135-170.

Araoz, A. (1993a), **A Report on Interviews at the WAITRO Seminar on "R&D Strategies for the 1990s" in Quebec, 13-15, October, 1992**, Unpublished mimeo.

Araoz, A. (1993b), **Restructuring of Science and Technology and New Modalities of Industrial Research in New Zealand**, A paper prepared for the UNIDO-IDRC Project on Industrial Technology Research Institutes.

Argyris, C. (1968), "On the Effectiveness of Research and Development," **American Scientist**, Vol. 56, Winter, pp. 344-355.

Armentrout, D.R. (1986), "Engineering Productivity Management and Performance Measurement," **Journal of Management Engineering**, Vol. 2, No. 3, pp. 141-147.

Asher, D.T. (1962), "A Linear Programming Model for the Allocation of R&D Efforts," **IRE Transactions on Engineering Management**, Vol. EM-9, No. 4, pp. 154-157.

Baker, N.R. (1974), "R&D Project Selection Models: An Assessment," **IEEE Transactions on Engineering Management**, Vol. EM-21, No. 4, pp. 165-171.

Baker, N.R. and Pound, W.H. (1964), "R&D Project Selection: Where We Stand," **IEEE Transactions on Engineering Management,** Vol. EM-11, No. 4, pp. 124-134.

Baker, N.R., and Green, S.G. and Bean, A.S. (1986), "Why R&D Projects Succeed or Fail," **Research Management**, Vol. 29, No. 6, pp. 29-34.

Baker, N.R., Siegman, J., and Rubenstein, A. (1967) "The Effects of Perceived Needs and Means on the Generation of Ideas for Industrial R and D Projects," **IEEE Transactions on Engineering Management**, Vol. 14.

Becker, R.H. (1980), "Project Selection Checklists for Research, Product Development and Process Development," **Research Management**, Vol. 23, No. 5, pp. 34-36.

Bedrosian, K. (1971), "Evaluating Your R&D Effort," **Innovation**, No. 22, June, pp. 46-55.

Beer, M. (1988), "Performance Appraisal: Dilemmas and Possibilities," in **Managing Professionals in Innovative Organizations**, R. Katz (ed.), Ballinger, Cambridge, MA, pp. 539-550.

Booker, J.M. and Bryson, M.C. (1985), "Decision Analysis in Project Management: An Overview," **IEEE Transactions on Engineering Management**, Vol. EM-32, No. 1, pp. 3-9.

Bowers, L.J. and Elliott, R.W. (1992), "Evaluation of Forty-four University Research Programs that Receive Line-Item Appropriations from the Texas Legislature," **SRA Journal**, Vol. 24, No. 1, pp. 57-65.

Bowman, C.W. (1992), "Measuring Good Research Management," **Research-Technology Management**, Vol. 35, No. 2, pp. 13-15.

Bozeman, B. and Loveless, S. (1987), "Sector Context and Productivity: A Comparison of Government and Industry R&D Units," **Administration & Society**, Vol. 19, No. 2, pp. 197-235.

Brandenburg, R.G. (1966), "Project Selection in Industrial R&D: Problems and Decision Processes," in **Research Program Effectiveness**, M.C. Yovits (ed.), Gordon and Breach.

Brandenburg, R.G. and Langenberg, F.C. (1969), "R&D Project Selection and Control at Crucible Steel Corporation," **Research Management**, Vol. 12, No. 2, pp. 123-139.

Bra Haupt, N. (1992), **Federal Laboratories and Competitiveness: An Evaluation Model,** Treasury Board, Office of the Comptroller General of Canada, Evaluation and Audit Branch, Ottawa.

Buchanan, B. (1974), "Government Managers, Business Executives, and Organizational Commitment," **Public Administration Review**, Vol. 34, pp. 339-347.

Burgess, J.S. (1966), "The Evaluation of a Government-Sponsored Research and Development Program," **IEEE Transactions on Engineering Management**, Vol. EM-13, No. 2, pp. 84-90.

Cardus, D., Fuhrer, M.J., Martin, A.W. and Thrall, R.M. (1982), "Use of Benefit-Cost Analysis in the Peer Review of Proposed Research," **Management Science**, Vol. 28, No. 4, pp. 439-445.

Cetron, M.J., Martino, J. and Roepcke, L. (1967), "The Selection of R&D Program Content - Survey of Quantitative Methods," **IEEE Transactions on Engineering Management**, Vol. EM-14, No. 1, pp. 4-13.

Chan, J.L. (1978), "Organizational Consensus Regarding the Relative Importance of Research Output Indicators," **The Accounting Review**, Vol. 53, pp. 309-323.

Christensen, H.C. (1987), "Integrated Internal Evaluation, in Ormala," in **Evaluation of Technical Research and Development: Experience of Nordic Countries**, E. Ormala (ed.), Nordic Cooperative Organization for Applied Research, Helsinki, pp. 122-140.

Chubin, D. E. (1987), "Research Evaluation and the Generation of Big Science Policy," **Knowledge**, Vol. 9, pp. 254-277.

Clarke, T.E. (1981), "R&D Management Bibliography 1981" (Research Abstract), **IEEE Transactions on Engineering Management**, Vol. EM-29, pp. 104-105.

Clarke, T. E. and Reavley, J. (1988), "Problems Faced by R&D Managers in Canadian Federal Government Laboratories," **R&D Management**, Vol. 18 No. 1, pp. 33-44.

Clogston, A. M. (1982), "Performance Evaluation of Research at Bell Laboratories," in **The Quality of Research in Science**, National Academy Press, Washington, DC, pp. 56-60.

Collier, D. W. (1977), "Measuring the Performance of R&D Departments," **Research Management**, Vol. 20, No. 2, pp. 30-34.

Collier, D. W. and Gee, R.E. (1973), "A Simple Approach to Post-Evaluation of Research," **Research Management**, Vol. 16, No. 3, pp. 12-17.

Connolly, T. (1972), **The Diffuse Decision: An Integrating Focus**, invited paper at Joint TIMS/ORSA/AIEE Meeting, Atlantic City, NJ, Nov., 1972.

Cordero, R. (1990), "The Measurement of Innovation Performance in the Firm: An Overview," **Research Policy**, Vol. 19, No. 2, pp. 185-192.

Cramer, R. H. and Smith, B. E. (1964), "Decision Models for the Selection of Research Projects," **The Engineering Economist**, Vol. 9, No. 2, pp. 1-20.

Crow, M. and Bozeman, B. (1987), "R&D Laboratory Classification and Public Policy: The Effects of Environmental Context on Laboratory Behavior," **Research Policy**, Vol. 16, No. 5, pp. 229-258.

Cutler, W. G. (1979), "R&D Planning - Formulating the Annual Research Program at Whirlpool," **Research Management**, Vol. 22, No. 1, pp. 23-26.

Danila, N. V. (1985), "Evaluating Pharmaceutical R&D," in **Managing High Technology: An Interdisciplinary Perspective**, B.W. Mar, W.T. Newell and B.O. Saxbergs (eds.), Elsevier Science Publishers, New York, pp. 25-40.

Dean, B. V. (1967), "A Research Laboratory Performance Model," **IEEE Transactions on Engineering Management**, Vol. EM-14, No. 1, pp. 44-46.

DeCicco, R. W. (1968), "Economic Evaluation of Research Projects - By Computer," **Chemical Engineering**, Vol. 75, No. 12, pp. 84-90.

De la Mothe, J. (1992), "The Political Nature of Science and Indicators," **Science and Public Policy**, Vol. 19, No. 6, pp. 401-406.

Doz, Y., Angelmar, R. and Prahalad, C. K. (1985), "Technological Innovation and Interdependence: A Challenge for the Large, Complex Firms," **Technology in Society**, Vol. 7, Nos. 2/3, pp. 105-125.

Englisch, H. and Czerwon, H.-J. (1990) "Quantification of the Performance of Research Units: A Simple Mathematical Model," **Research Policy**, Vol. 19, pp. 477-480.

FEICRO (1993) **Cooperative Research in European Perspective: Synthesis Report**, Commision of the European Communities, DG 13 C/4, SPRINT programme, Brussels.

Freeman, C. (1969), **Measurement of Output of Research and Experimental Development**, UNESCO, Paris.

Fung, D. and Shapiro, J. (1976), "Selection and Evaluation of Research Projects," **Chemistry in Canada**, Vol. 28, No. 2, pp. 19 & 21.

Galloway, E.C. (1971), "Evaluating R&D Performance - Keep it Simple," **Research Management**, Vol. 14, No. 2, pp. 50-58.

Gear, A.E., Lockett, A.G., and Pearson, A.W. (1971), "Analysis of Some Portfolio Selection Models for R&D," **IEEE Transactions on Engineering Management**, Vol. EM-18, No. 2, pp. 66-76.

Gerloff, E.A. (1973), "Performance Control in Government R&D Projects: The Measurable Effects of Performing Required Management and Engineering Techniques," **IEEE Transactions on Engineering Management**, Vol. EM-20, No. 1, pp. 6-13.

Gibbons, M. (1984) "Methods for the Evaluation of Research," in **Programme on Institutional Management in Higher Education, Seventeenth Special Topic Workshop, Evaluation of Research and Resource Allocation**, OECD, December 3-5.

Glass, E.M. (1970) "Evaluating R&D Organizations," **Research Development**, Vol. 21, No. 1, pp. 23-27.

Gold, B. (1989), "Some Key Problems in Evaluating R&D Performance," **Journal of Engineering and Technology Management**, Vol. 6, No. 1, pp. 59-70.

Government of Australia (1977), **Report of the Independent Inquiry into the Commonwealth Scientific and Industrial Organization, August 1977**, Australian Government Publishing Service.

Guba, E.G. and Lincoln, Y.S. (1989) **Fourth Generation Evaluation**, Sage Publications, Newberry Park, CA.

Hackett, J.W. (1962), "Evaluation of Research," **Research Management**, Vol. 5, No. 3, pp. 177-192

IDRC (1993), **Institutional Evaluation: A Framework for Building Organizational Capacity for IDRC's Research Partners**, draft, Evaluation Unit, Corporate Affairs and Initiatives Division (prepared by Universalia Management Group), International Development Research Centre, Ottawa, Oct.

Jain, R.K. and Triandis, H.C. (1990), **Management of Research and Development Organizations: Managing the Unmanageable**, John Wiley & Sons, Toronto.

Kelley, R. and Caplan, J. (1993), "How Bell Labs Creates Star Performers," **Harvard Business Review**, Vol. 71, No. 4, pp. 128-139.

Krawiec, F. (1984), "Evaluating and Selecting Research Projects by Scoring," **Research Management**, Vol. 27, No. 2, pp. 21-25.

Liberatore, M.J. and Titus, G.J. (1983), "The Practice of Management Science in R&D Project Management," **Management Science**, Vol. 29, No. 8, pp. 962-974.

Lucas Jr., R.E. (1971), "Optimal Management of an R&D Project," **Management Science**, Vol. 17, No. 11, pp. 679-697.

Lukkonen-Gronow, T. (1987), "Scientific Research Evaluation: A Review of Methods and Various Contexts of Their Application," **R&D Management**, Vol. 17, No. 3, pp. 207-221.

Mansfield, E. (1968), **Econometric Studies of Industrial Research and Technological Innovation**, New York: W.W. Norton.

Mansfield, E. and Brandenburg, R. (1966), "The Allocation, Characteristics and Outcome of the Firm's Research and Development Portfolio: A Case Study," **The Journal of Business**, Vol. 39, No. 4, pp. 447-464.

Mansfield, E. and Wagner, S. (1975), "Organizational and Strategic Factors Associated with Probabilities of Success in Industrial R&D," **The Journal of Business**, Vol. 48, No. 2, pp. 179-198.

Mansfield, E. et al., (1971) **Research and Innovation in the Modern Corporation**, W.W. Norton & Co., New York.

Marquis, D.G. (1969), "The Anatomy of Successful Innovations," **Innovation**, No. 7, pp. 28-36.

Martin, B. R., Minchin, N. et al. (1987), "The Assessment of Scientific Research: A Case-Study of CERN," **Interdisciplinary Science Reviews**, Vol.12 No.1, pp.70-76.

Miller, R. (1992) "The Influence of Primary Task on R&D Laboratory Evaluation: A Comparative Bibliometric Analysis," **R&D Management**, Vol. 22, No. 1, pp. 1-20.

Montgomery, H. and Hemling, S. (1987), "Conceptions of Scientific Quality" in **Evaluation of Research: Nordic Experience**, FPR Publ., No. 5, Nordic Science Policy Council, Copenhagen, pp.17-126.

Mudola, D. M. (1983), "The Pathology of Institution Building: The Tanzanian Case," in **Challenging Rural Poverty**, R.G. Kiros (ed.), Organization for Social Science Research in Eastern Africa, Congress (1), Alemaya, Ethiopia.

Mullen, J. (1992) **Some Recent Canadian Initiatives in the Deveopment of Industrial Technology**, Ottawa, mimeo, Feb.

Neufeld, G. A. (1992), **A Think-Piece Concerning the Evaluation of Science & Technology Programs**, Bronson Consulting Group, Ottawa, 8 May.

NRC (1992) **National Research Council Operational Plan Framework**, unpublished mimeo.

OECD (1987a), **Government Research Establishments: Synthesis Report**, Directorate for Science and Technology, Organisation for Economic Co-operation and Development, Paris, October 1987.

OECD (1987b), **Evaluation of Research: A Selection of Current Practices**, OECD, Paris.

Oldham, G. (1988), **Commentary on the ASTEC Review of CSIRO**, Science Policy Research Unit, University of Sussex, Occasional Paper No. 6, December.

Pelz, D.D. and Andrews, F.M. (1966) **Scientists in Organizations**, John Wiley and Sons, New York.

Persson, O. (1987), "Bibliographies and bibliometrics - Some Problems of Measurement" in: **Evaluation of Research: Nordic Experience**, FPR Publ. No.5, Nordic Science Policy Council, Copenhagen, pp. 186-202.

Porter Jr., J.G. (1978), "Post Audits - An Aid to Research Planning," **Research Management**, Vol. 21, No. 1, Jan., pp. 28-30.

Ritchie, E. (1970), "Research on Research: Where Do We Stand," **R&D Management**, Vol. 1, No. 1, pp. 3-9.

Roberts, C.S. (1959), "Product Selection - Witchcraft or Wisdom," **IRE Transactions on Engineering Management**, Vol. EM-6, No. 3, pp. 68-71,

RockCliffe Research (1992), **Models and Policies for the Commercialization of Government Science and Technology**, RockCliffe Research and Technology Inc., Ottawa, 21 Aug.

Sanchez, A.M. (1969), "R&D Project Selection Strategy: An Empirical Study," **R&D Management**, Vol. 19, No. 1, Jan., pp. 63-68.

Schainblatt, A.H. (1981), **Measuring the Productivity of Scientists and Engineers in R&D: A State of the Practice Review**, The Urban Institute, Washington, May.

Schroeder, H.H. (1971), "R&D Project Evaluation and Selection Models for Development: A Survey of the State of the Art," **Socio-Economic Planning Sciences**, Vol. 5, No. 1, pp. 25-39.

Shankar, P.S. (1985) "Evaluating the Evaluation of Research and Development", **Journal of Scientific and Industrial Research**, Vol. 44, May, pp. 225-229.

Smith, Keith (1992), "Technological Innovation Indicators: Experience and Prospects," **Science and Public Policy**, Vol. 19, No. 6, pp. 383-392.

Smith, William E., et al. (1980), **The Design of Organizations for Rural Development Projects**, IDII (Organizing for Development, an International Institute), World Bank Staff Working Paper No. 375, Washington, D.C., March.

Stolte-Heiskanen (1986) "Evaluation of Scientific Performance on the Periphery, **Science and Public Policy**, April, pp. 83-88.

WAITRO (1972a), **International Seminar on Interaction of R&D Institutes With Local Industry: Final Report**, World Association of Industrial and Technological Research Organizations, Publication No. 3.

WAITRO (1972b), **Priority of Needs of Industrial Research Institutes in Developing Countries**, World Association of Industrial and Technological Research Organizations, Publication No. 3.

Wallmark, J.T., McQueen, D.H., and Sedig, K.G. (1988), "Measurement of Output from University Research: A Case Study," **IEEE Transactions on Engineering Management**, Vol. 35, No. 3, pp. 175-180.

Yencken, D. (1987), **Ministers and Masters: Public Policy and University Research**, University of Melbourne Faculty of Architecture and Planning, July 1987.

Ziman, J. (1987), ***Science in a "Steady State": The Research System in Transition***, Science Policy Support Group Concept Paper No. 1, C/o AFRC, London.