Pergamon

0305-0483(95)00006-2

Omega, Int. J. Mgmt Sci. Vol. 23, No. 3, pp. 281–294, 1995

0305-0483/95 $9.50 + 0.00

An Integrated Cost-performance Model of Research and Development Evaluation

E GEISLER

University of Wisconsin-Whitewater, USA

(Received January 1994; accepted after revision January 1995)

This paper proposes an integrated model of R & D evaluation which links the cost of research with its various outputs, from proximal or immediate to the intermediate outputs in the client organizations of the R & D performer. This model differs from previous models in that it consolidates both cost and performance assessment. Indexes of cost-performance are constructed, for each category of output, and as a measure of productivity per scientist and engineer engaged in research. The indexes proposed in this model reflect the relative cost-effectiveness of the R & D activity throughout a substantial portion of the innovation process. As a comprehensive and integrated model, it transcends the limitations of immediate outputs utilized by other models. The model allows for comparisons, at the program, corporate and national levels, of the cost-performance of research. Some methodological and implementation problems are discussed.

Key words—research, research and development, R & D, management of science/technology, performance ratios, evaluation

1. INTRODUCTION

THE EVALUATION OF THE impacts of research has long been the topic of a continuous effort by academics and practitioners [6, 35]. The importance of an adequate evaluation is magnified in light of the growing consensus that innovation and technology leadership are paramount in the comparative advantage of companies and nations [16, 17, 23].

This paper proposes an integrated cost-performance model of research and development (R & D) evaluation. The model incorporates inputs to and outputs from the research activity. The model proposed here differs from previous models in that it consolidates two distinct models: cost and performance assessment [13, 14]. Each of these models was designed to allow for programmatic as well as institutional and national assessments of R & D. The integrated cost-performance model proposed here is also suited for assessment of R & D at different levels of the organization.

The need for integration of cost and performance models of research evaluation is evident in comments made by several writers [5, 25, 36, 42]. There are three key factors which seem responsible for the gap: complexity of the phenomenon; disciplinary boundaries of economics and behavioral approaches to research management and evaluation; and problems of imputation across stages in the lengthy process of the research activity. The integrated model of cost-performance proposes to bridge the gap and to offer a workable framework for evaluation.

Research is defined in the traditional classification as basic and applied research. The National Science Board defines basic research as: "... research that advances scientific knowledge, but does not have specific commercial objectives". Applied research is: "... directed toward gaining knowledge or understanding necessary for determining the means by which a recognized and specific need may be met" [41, p. 94]. Development is defined as "the systematic use of the knowledge or understanding gained from research directed toward the production of useful materials, devices, systems, or methods, including the design and development of prototypes and processes" [40, p. 94].

Evaluation is defined in this paper as an activity directed toward establishing the value of research and development to the organization in which it is undertaken for the assessment of other activities (which benefit from R & D), and for improved allocation of resources to R & D. Evaluation includes establishing value of the *direct* as well as *indirect* impacts of R & D, through all stages of the downstream flow of the R & D/innovation process [3, 15, 20, 23]. In this study the downstream flow is divided into four stages and four output categories as proposed by Rubenstein and Geisler [30, 36] and Geisler [14]. The four stages are: (1) the R & D process itself; (2) transforming and diffusing organizations; (3) social/economic sub-systems; and (4) the Society and the Economy. The output categories are: (1) immediate outputs; (2) intermediate outputs; (3) pre-ultimate outputs; and (4) ultimate outputs. The categories of outputs correspond to the four stages.

2. REVIEW OF PREVIOUS MODELS

The literature on R & D management may be classified into four major categories of studies on R & D evaluation. The first includes models assessing the economic impacts of research. Papers by Mansfield [23, 24] and Griliches [16] provide examples of linkage between inputs to research and the economic impacts downstream on the firm's innovation effort. Another example is the work of McGrath and Romeri [21], who proposed a R & D Effectiveness Index which compares profits from new products with investments in R & D that presumably was instrumental in these new products. The second category includes models of research performance in terms of the productivity of individual researchers and their groups [4, 5]. The third category includes models of valuation of research measured by selected outcome indicators, such as counts of publications, citations and patents. Irving *et al.* [18] and Martin *et al.* [25] assessed specific disciplines and trends in national research. Tijssen [39] assessed energy research, whereas Narin [29], Sen and Shailendra [38], and Trajtenberg [40] utilized various counts of output indicators to assess research impacts on the organization and its innovation effort. Finally, the fourth category includes models of evaluation which employ subjective assessments, primarily in the form of peer review by individuals or panels of experts [2, 3, 7, 34].

These four categories of models may be further classified into two types: (1) input (cost) models, and (2) output (performance) models. The input models of evaluation compare the levels, distribution and intensity of investments in R & D [6, 10, 19]. They allow for comparison among R & D organizations as to how much should be invested in R & D, and the construction of R & D-intensity indexes and benchmarks of the ratios of R & D investments to other economic indexes such as sales, profits and revenues [17, 21, 32, 41]. The input models do not assess performance of R & D, but the data they utilize are also used in the economic outputs model listed above.

The economic outputs model is the only category of the four listed above that relates inputs to and outputs from R & D in an attempt to assess the benefits derived from R & D. These models have methodological problems of two kinds: (1) isolation of the economic/financial benefits of R & D to the organization from other effects, such as efficiencies in manufacturing, management and marketing; and (2) imputation problems of the time lag between R & D and economic benefits to the organization. These models make strong assumptions to allow for a causal link between inputs and outputs. They include a 'leap of faith' between inputs and outputs. Much of the uncertainty and the gap resides in the research activity and its link to measurable benefits, and less in the development portion of the innovation process.

The integrated model proposed in this paper is an attempt to overcome some of the shortcomings of previous models. In the case of input models, the lack of the link to outputs will be filled through integration of cost and performance. In the case of output models, the problems of economic models of isolation and imputation will be resolved through the stage approach to performance assessment. Instead of making a 'leap of faith' from inputs to benefits, the proposed integrated model traces forward the outputs from R & D through the stages and transformations in the R & D-innovation process. In the case of output models employing various indicators, the proposed model integrates these and other indicators derived empirically in R & D organizations.

3. DEVELOPING COST AND PERFORMANCE MODELS

The proposed integrated model of R & D evaluation incorporates a cost and a performance model previously developed to independently assess inputs to and outputs from R & D.

3.1. Cost model

A cost model is essentially an input model in which the inputs to R & D are calculated in the form of indexes of direct investments in R & D and their ratios to other inputs and/or outputs. Cost items in the model include expenditures on scientists and engineers; support personnel; materials and supplies; and plant and equipment [41]. In the case of research programs, indirect costs (overhead) may be added. Organizational assessment of indirect costs are also additions to total expenditures for R & D—following the organization's own criteria for allocation of such costs [19, 27, 31].

Geisler [13] proposed the development of indexes of *cost per article*, *cost per patent*, and *cost per scientist and engineer* (S & E). These indexes assume all expenditures for *research* divided by a category of input (S & E), and selected categories of outputs (articles and patents). The indexes allow for a comparison over time among nations and industries. Comparing among companies would be more difficult in view of poor availability of reliable data [28, 42]. Expenditures for *development* should be added since patents are usually considered outputs from the development portion of the innovation process. Other indicators of outputs from development include, for example, new products conceived and key improvements, introduced in products or services.

However, for the integrated model it is sufficient to have a figure of total cost of research and development and the cost per scientist and engineer. In the case of total cost of R & D the model is designed to calculate aggregated costs at each progressive stage in the R & D/innovation flow [35]. In this way the distinction between research and development is resolved since the activities culminate in downstream outputs that are captured by the model. The cost per scientist and engineer is also cumulative at each stage and can be adjusted to reflect unit or program boundaries (e.g. project, department), and temporal aggregation along the stages of the innovation flow (e.g. cost of all S & Es who had contributed to the new product up to the stage of initial prototype).

The indexes in this cost model will be incorporated into the integrated model of R & D evaluation. They are joined by the indexes of output indicators in the model of evaluation of R & D performance.

3.2. Performance model

This model is based on the development of key or leading output indicators for each of the stages in the downstream process of R & D impact assessment [7, 8, 14, 36]. The model has four stages or categories of outputs: (1) *immediate* or direct/proximal outputs such as publications and patents; (2) *intermediate* outputs, which are outputs of the transforming organizations such as companies, and inputs to social and economic sub-systems, such as new products, new materials and new methods; (3) *pre-ultimate* outputs, which are products and services generated by the economic and social subsystems; and (4) *ultimate* outputs, which are those things of value to the economy and society.

The *immediate* outputs are the outcomes from the R & D activity itself, measured by indicators such as publications, citations, and patents. These outputs are then transferred to transforming organizations such as the business entities of the corporation, engineering, and manufacturing. These entities process and transform the immediate outputs they have absorbed. The outcomes of these transformations are the *intermediate* outputs, measured by indicators such as new products, new materials, new methods for analysis, reductions in cost and improved productivity. These are outputs internal to the organization. These two categories of outputs are usually used in assessments of the R & D function. Intermediate outputs are compared with inputs to R & D and with the immediate outputs, particularly in economic models.

However, the innovation flow includes two additional categories of outputs seldom used in R & D assessment [35, 36]. Pre-ultimate outputs are generated by the social and economic entities which then export them downstream to their environment. Such outputs are external to the organization. They are measured by such indicators as improved competitiveness, opening new markets, and finding a cure for a disease. Finally, these pre-ultimate outputs are absorbed by the economy and the society and produce the ultimate outputs, measured by such indicators as quality of life, economic growth, as well as negative effects such as increased pollution and other environmental concerns.

As elaborated below in the description of the integrated model, the stages in the R & D/innovation flow are temporal and conceptual. They represent a serial movement of outcomes along a time dimension, and are also measuring distinct phenomena. For example, immediate outputs assess the outcomes from the R & D activity itself, whereas intermediate outputs assess outcomes from the business (or non-R & D in not-for-profit organizations) side of the organization. In this latter case, the generation of the intermediate outputs is governed by a set of factors very different from those in the R & D phenomenon.

The movement from one stage of outputs to another is negotiated through diffusion mechanisms that are inherent in the innovation process and have received some attention in the literature [3, 11, 14, 28, 37]. The integrated model concentrates on relationships among output categories. It considers the mechanisms of diffusion and the factors hindering or facilitating the flow as explanatory variables needed for an in-depth analysis of each individual case of R & D/innovation flow and the host organization. Problems of imputation are not eliminated, but they are reduced by the model's tracking of the flow through the various stages or categories of outputs. This is primarily due to the fact that additional insights are obtained into the way the R & D/innovation flow occurs and assessed at progressive milestones, rather than in one comprehensive 'leap of faith' from inputs to R & D to intermediate or pre-ultimate outputs.

The model thus tracks the downstream contributions of R & D via sets of leading or key indicators. These indicators are roughly grouped in two categories: (1) *core* indicators, which are applicable across R & D organizations (generic indicators), and (2) *organization-specific*, which are unique to the organization performing the research.

In some respects there may be a cross-over between the core and organization-specific indicators, so that they are not mutually exclusive. It is assumed that the core indicators measure the phenomenon of the performance of the R & D organization. They were generated by the author and his colleagues from the R & D management literature as those indicators that, for each output category, would apply across different R & D performing organizations. As shown below, the integrated model was partially tested in two federal R & D laboratories, but the core indicators which were used in these laboratories also apply to R & D units in business corporations or to independent commercial R & D laboratories. The leading core output indicators for the immediate and intermediate stages are shown in Table 1.

3.2.1. Leading organization-specific output indicators. The organization-specific output indicators express some specific characteristics of the R & D organization being evaluated [14]. Their selection was by judgemental assessment in the two laboratories in which the exploratory study was conducted. A group of 12 R & D managers and 19 scientists/engineers were asked to list those measures of immediate and intermediate outputs from their laboratories which characterize their organization to the extent that they may differentiate their laboratory from other laboratories [45]. Ten illustrative measures thus obtained are shown in Table 2.

With the assistance of a sub-set of 2 R & D managers in one of the laboratories, the measures were then grouped into two categories each for the immediate and intermediate outputs, giving four leading indicators, as shown in Table 2. In agreement with some recent findings in the literature on R & D commercialization, the R & D managers listed those measures for intermediate outputs which measure the transforming organization's commitment to commercialize immediate outputs from R & D [1, 3, 8, 10, 12, 26, 37].

3.2.2. Development of index of leading indicators. The next step in the model of

Table 1. Leading core output indicators for the immediate and intermediate R & D output categories

I. IMMEDIATE OUTPUTS
1. Index of written scientific and technical outputs
Measures
1.1 Number of publications in refereed journals
1.2 Number of technical reports
1.3 Number of patents
1.4 Number of patent disclosures
1.5 Number of citations in refereed journals
2. Index of hardware/software/other outputs
Measures
2.1 Number of new products conceived
2.2 Number of key improvements suggested
2.3 Number of new and improved test methods, models, standards, concepts and databases transferred downstream
2.4 Number of new ideas transferred downstream
2.5 Number of problems solved for users/clients downstream
2.6 Improved understanding of phenomenon
3. Index of overall reputation of R & D performers
Measures
3.1 Number of complaints by clients/users
3.2 Judgment by clients/users
3.3 Judgment by other R & D performers (peers)
3.4 Number of awards received
3.5 Milestones/objectives met
II. INTERMEDIATE OUTPUTS
1. Index of scientific/technical impacts on direct user of R & D outcomes
Measures
1.1 Number of improved and/or new products
1.2 Number of improved and/or new materials
1.3 Number of improved and/or new tests and methods of analysis
2. Index of economic impacts on direct user of R & D outcomes
Measures
2.1 Actual cost reduction/savings in products or processes
2.2 Actual improvement in productivity of material/equipment/techniques or people
2.3 Actual improvements in performance in sales, profits, sophistication of new and/or improved products and services
3. Index of responsiveness of R & D
Measures
3.1 Judgment by the direct client/user downstream
3.2 Judgment by other (non-R & D) organizations
3.3 Judgment of other R & D performing organizations

performance evaluation of R & D provides a detailed elaboration of the procedure described in [14]. For each stage of outputs, a basic set of key leading indicators is weighted by degree of importance by the Weighted Attributes Method [14, 40, 43]. The indicators and their measures are selected from those in Tables 1 and 2. They are selected under a condition of orthogonality so as to avoid multiple counting. Furthermore, validation for independence is based on three criteria: (1) indicators, via their measures, describe as many different organizational characteristics as possible; (2) limited number of indicators used and formation of a small set of key indicators; and (3) resulting set containing mutually exclusive indicators. These criteria

Table 2. Leading organization-specific indicators for the immediate and intermediate R & D output categories

I. IMMEDIATE OUTPUTS
1. Level of technical expertise
Measures
1.1 Ratio of doctorate holders to scientific workforce
1.2 Relative experience of scientists and engineers (S & Es): total years of technical work
2. Attractiveness of R & D organization
Measures
2.1 Number of candidates applying for each position in the technical/scientific area
2.2 Age profile of S & Es
2.3 Judgment by other R & D performers (peers)
II. INTERMEDIATE OUTPUTS
1. Level of investment in exploitation of R & D outcomes
Measures
1.2 Annual funds allocated to technology commercialization
1.2 Number of personnel from non-R & D units working with R & D
2. Level of importance of R & D outcomes
Measures
2.1 Role of new products/services in the organization's success and survival (judgmental)
2.2 Perceived success (trade record) of outcomes transferred from R & D in the organization's performance (judgmental)
2.3 Judgment by other organizations (peers)

were utilized in the generation of both the core and organization-specific indicators.

As shown in Tables 1 and 2, each indicator is measured by a small set of measures. The measures assess the progress of the indicator on a time dimension by the differences in the measures between two time periods. For example, the indicator *index of written scientific and technical outputs* is measured by (among other things) the number of patent disclosures, as the difference between disclosures in period x (last year) and period $x+1$ (current year) [30, 33]. The value of the indicator is itself derived as a weighted combination of its measures.

In the R & D organization, the weights for measures and indicators will be assigned by the evaluators, who are those people responsible for assessment of the activity (R & D managers and/or outside consultants). In the example in Appendices A and B, a sample of 12 R & D managers and 19 bench scientists provided the relative weights.

Positive normalized weights (w_{ia}) are first assigned to the measures of each leading core and organization-specific indicator (or simply, indicator), so that:

$$0 < w_{ia} < 1 \quad a = 1, 2, \ldots, n(i), \tag{1}$$

and

$$\sum_{a=1}^{n(i)} w_{ia} = 1 \tag{2}$$

where i = the number of the indicator and $n(i)$ = number of measures of the ith indicator. For each indicator the value is the sum of the weighted measures, so that the index value for the ith indicator is

$$IV_i = \sum_{a=1}^{n(i)} d_{ia} w_{ia} \tag{3}$$

where

w_{ia} = weight of the ath measure of indicator i
$n(i)$ = number of measures of the ith indicator.
d_{ia} = values of ath measure of indicator i.

We now proceed to obtain the Indexes of Leading Output Indicators (LOIs) corresponding to each stage of output. A similar weighting procedure is used on the index values for the indicators, with a new set of weights (W_{is}) so that:

$$\text{LOI}_s = \sum_{i=1}^{n} IV_i W_{is} \tag{4}$$

in which s = stage of output (s = 1–4), i.e. immediate, intermediate, preultimate and ultimate, and n = number of leading output indicators. ($W_{is} = 0$ if indicator i does not bear on stage s.)

3.2.3. Development of index of key outputs indicators. The LOIs described above assess the outputs from R & D in each of the four stages along the R & D innovation continuum. However, for a more comprehensive evaluation it is necessary to aggregate the LOIs for any stage. Thus, evaluation at the stage of intermediate outputs would require aggregation of LOI_1 and LOI_2.

The result is an overall Key Outputs Indicators (KOI) value as a figure of merit of R & D. This index is now obtained by a new set of normalized weights of the LOI_s so that

$$\text{KOI} = \sum_{s=1}^{4} \text{LOI}_s \hat{W}_s \tag{5}$$

in which

s = stage of output
$\hat{W}_s$ = normalized weights for each LOI.

As in the procedure for LOI, weights are assigned by evaluators in the organization and indicate the importance or emphasis on the proximity of the value to the R & D activity. For example, emphasis on immediate outputs measures the degree of importance management places on short term and/or tangible outcomes from R & D.

In summary, there are three computations in this procedure. All three utilize the same method of multiplying values by normalized weights, albeit a different set of weights for each computation. The first computes the value for each LOI (IV_i) by the sum of the values for the measures for the indicator times their normalized weights W_{ia}. The second computes the Index of LOIs by the sum of the values of the leading output indicators times their normalized weights W_{is}. Finally, we compute the overall index of KOI by the sum of the (one to four) LOIs times their normalized weights $\hat{W}_s$.

Appendix A gives an illustration of a study in the two aforementioned federal R & D laboratories. The Appendix shows the development of the indexes and the evaluative comparisons made between the two organizations.

4. AN INTEGRATED MODEL OF COST-PERFORMANCE

4.1. Development of cost-performance index

The cost-performance model manipulates cost and performance indexes for each of the four stages of R & D outputs to form ratios of performance per cost, so that:

Index of Cost Performance (ICP)

$$= \frac{\text{LOI}_s}{\text{Cost of R \& D (total expenditures)}} \quad (6)$$

where

LOI_s = Index of Leading Output Indicators
s = stages of outputs from R & D ($s = 1$–4).

The LOI_s is an evaluative index expressing a 'moment-in-time' or a 'snapshot' of where the organization stands in its R & D outputs, as well as how the outputs from R & D have progressed along the innovation continuum. Therefore the ICP also provides a 'snapshot' evaluation by organization, department, or project—as long as cost may be associated with the unit and allocated to it. The cost of R & D would thus be the cumulative expenditures which can be identified in the R & D organization (for immediate outputs), and in all organizations involved (for intermediate outputs and beyond, downstream).

The ICP for the immediate outputs, for example, expresses the relative effectiveness of producing immediate outputs from R & D—per investments in their production [13, 31, 22]. The cost includes human resources, capital investments, materials, supplies, communications and overhead [1, 2, 6, 13].

The ICP for KOI_2 will be the weighted sum of LOI_1 and LOI_2 (immediate and intermediate outputs), per the total cumulative cost (actual expenditures) in both the R & D producing organization and the transforming organization. Therefore, this cost entails: (1) cost of producing the R & D outputs, (2) cost of transfer and diffusion, and (3) cost of integrity R & D outputs into the transforming organization's efforts to generate its own outputs (e.g. products, processes, services).

A similar index would be ICP per cost of scientist and engineer (S & E), so that:

Index of cost/performance per scientist & engineer

$$= \frac{\text{LOI}_s}{\text{Cost per S \& E}} \quad (7)$$

(total cumulative expenditures for R & D divided by number of S & E engaged in R & D).

This index expresses the relative productivity of the scientific personnel engaged in R & D, as measured by the outputs from their activity. When this index is constructed for the same time period (e.g. budget-year) for inter-unit comparisons, it *will* reflect the relationship between cost and outputs.

However, when this index shows differences over time for a given R & D and/or downstream unit, such differences may be partially explained by a lag between expenses and output. Therefore, in the case of time-series data, the LOI has to be reconstructed with the addition of an adequate and individualized time lag for *each indicator*. For example, written scientific outputs may have a lag of 1–2 years from the time R & D resources are expended. Overall reputation may take 3–5 years to be generated. The different time-lags may also account for differences in uncertainties associated with the generation of the various outputs.

4.2. Development of index of overall cost/performance

The index of overall KOIs as shown in equation (5) is the aggregate weighted sum of LOI_1 through LOI_4, for all the four stages of downstream outputs. The KOI may be compared with costs, so that

Overall Index of Cost/Performance of R & D

$$= \frac{\text{KOI}}{\text{Total Cost of Research}} \quad (8)$$

(total cumulative expenditures for R & D/innovation aggregated over the whole time period, from initial R & D to impacts on the economy and society).

The computation of total cumulative cost of R & D in equation (8) may have a lag of several, perhaps many years from the time expenditures for R & D have occurred till the impacts of the outputs on the economy and society become measurable. For example, *pre-ultimate* outputs may be: improvements in level of mortality; level of morbidity; improvements in safety of work environment; improved mobility; and reduction or extinction of particular causes of death. *Ultimate* outputs may be: energy independence; national security, quality of life; and improved GNP.

Therefore, for reasons of feasibility of use and availability of data, in the examples in Appendices A and B only the immediate and intermediate outputs have been assessed. Appendix B provides an illustration of the indexes of cost/performance of R & D for the stage of immediate and intermediate outputs. Although it is very difficult to connect pre-ultimate and especially ultimate outputs to individual R & D efforts, it may be feasible by tracing forward the downstream impacts of R & D in a manner akin to the historical method of the Hindsight Project [9]. Clearly, the impacts become much more diffused and difficult to trace and to isolate as we move downstream the R & D/innovation flow [43, 44].

However, in the cases of the federal R & D laboratories referred to in Appendices A and B, some of the pre-ultimate and ultimate outputs are inherent in the mission of the laboratories because of the national nature of these organizations. This is further elaborated in the Appendices.

There are also costs associated with the 'movement' of the immediate outputs of R & D (e.g. publications, patents, new ideas) through the various stages of the downstream flow of innovation. These are costs incurred, for instance, by corporations when they adapt and transform the immediate outputs into new and improved products, processes, and services, which in turn contribute to the pre-ultimate and ultimate outputs. These costs are *not* costs of R & D, although they do represent an integral component of the total cost of R & D impacts [9, 30, 32].

Appendix C shows an example of how indicators of pre-ultimate and ultimate outputs may be linked conceptually to individual R & D effort along the R & D/innovation flow.

This overall index of cost/performance of research is qualified by several factors discussed above, such as time lag, distance between stages in the phenomenon of R & D/innovation, and the assignment of costs along the R & D/innovation continuum [6, 12, 25, 31]. Nevertheless, this overall index offers an approximation to the relation between the aggregated impacts of R & D on downstream entities in the economy and society, and the costs of the R & D activity which generated these impacts. Clearly, R & D is an ongoing and integrative activity, and measures of cost and budgetary timeframes are artificial cuts in a continuous effort. However, the procedures described in this paper to generate the KOI and the overall index of cost/performance of R & D will allow for an improved measurement of the relation between R & D/innovation and its impacts downstream [35, 36].

5. CONCLUSIONS

This paper proposed the generation of indexes of cost-performance of R & D by employing four stages/categories of outputs along the downstream flow of the innovation process. Issues of methodology, imputation, time lags and costs-assignments were addressed.

The indexes by total cost of R & D and by scientist and engineer are useful tools for evaluation of R & D at the program, corporate, and national levels. At the program and corporate levels these indexes allow for assessment of research beyond the immediate outputs of publications and patents. The indexes provide a mechanism for a company to asses the impacts of its research on its products, services, processes, and its clients, and to compare these impacts with the costs it incurred in the research activity, and per scientist and engineer. The indexes allow for intercompany comparisons, as well as comparisons over time for the same program or company.

There are various methodological and phenomenological problems yet to be resolved in the construct of these indexes [45]. In summary, the indexes and procedures in this paper offer a comprehensive and integrated model which may transcend the limitations of the immediacy of research outputs, and extends the indexes to the downstream innovation process.

APPENDIX A

Illustrative Development of Indexes of Leading Output Indicators for Immediate and Intermediate Outputs for Two R & D Laboratories[1]

(1) Immediate outputs

The core indicators, their measures and weights are shown in the table below. In addition, the benchmarks of superior performance are also provided.

Laboratory A
Core Indicators

Indicator	Measure	Time period[2]	Weight[3]	Benchmark[4]
1. Scientific & technical outputs	20 Publications in refereed journals	9/92–9/93	35	30
	15 Patent applications	9/92–9/93	65	10
2. Hardware/software/ other outputs	12 new/improved test methods	9/92–9/93	60	5
	suggested 3 standards developed	9/92–9/93	40	5

Organization-Specific Indicators

Indicator	Measures	Time period	Weight	Benchmark
1. Technical expertise	10% doctorate holders	9/1/93	50	30%
	6 years average experience of S & Es	9/1/93	50	5 yr
2. Attractiveness of organization	3 candidates on average applying for S & E positions	9/92–9/93	70	5 candidates
	[200 miles] the proximity to nearest university [1][5]	9/1/93	30[6]	20 miles

From the formula in equation (3):	Weights assigned to IV_i:
$IV_1 = 16.75$ (Benchmark = 17.00)	$IV_1 = 0.4$
$IV_2 = 8.40$ (Benchmark = 5.00)	$IV_2 = 0.2$
$IV_3 = 8.00$ (Benchmark = 17.50)	$IV_3 = 0.2$
$IV_4 = 0.30$[6] (Benchmark = 0.90)	$IV_4 = 0.2$

Applying equation (4):

$$\mathrm{LOI(A)}_1 = 10.04 \quad [\text{Benchmark} = 11.48]$$

Laboratory B:
Core Indicators

Indicator	Measure	Time period	Weight	Benchmark
1. Scientific & techncial outputs	43 publications in refereed jouranals	9/92–9/93	75	30
	7 patent applications	9/92–9/93	25	5
2. Hardware/software/ other outputs	2 new/improved test methods	9/92–9/93	10	0
	about 20 *useable* ideas/ concepts transferred downstream[7]	9/92–9/93	90	20

Organization-Specific Indicators

Indicator	Measure	Time period	Weight	Benchmark
1. Technical expertise	30% doctorate holders	9/1/93	90	30%
	8 yr average experience of S & Es	9/1/93	10	5 yr
2. Attractiveness of organization	6 candidates on average applying for S & E positions	9/92–9/93	20	5
	10 miles from nearest[6] university	9/1/93	80[6]	20

	Weights assigned to IV_i:
$IV_1 = 34.0$ (Benchmark = 23.75)	$IV_1 = 0.6$
$IV_2 = 18.2$ (Benchmark = 18.00)	$IV_2 = 0.1$
$IV_3 = 27.8$ (Benchmark = 27.50)	$IV_3 = 0.2$
$IV_4 = 2.40$ (Benchmark = 2.40)	$IV_4 = 0.1$

Applying equation (4):

$$\mathrm{LOI(B)}_1 = 28.02 \quad [\text{Benchmark} = 21.75].$$

(2) Intermediate outputs[8]

Laboratory A[9]

Indicator[13]	Measure	Time period[10]	Weight	Benchmark[11]
1. Scientific technical impact on direct user of R & D outcomes	2 new/improved products	9/93	60	2
	improved tests	9/93	40	2
2. Economic impacts on direct users of R & D outcomes	\$10 K actual cost reduction in manufacture of a product	9/92–9/93	90	NA[12]
	\$20 K actual improvement in a technique for testing materials	9/92–9/93	10	NA

By applying equation (3):	Weights assigned to IV_i:
$IV_1 = 2.4$ (Benchmark = 2.0)	$IV_1 = 0.3$
$IV_2 = 11.0$ (Benchmark = NA)	$IV_2 = 0.7$

$$\mathrm{LOI(A)}_2 = 8.42 \quad [\text{Benchmark} = \text{NA}]$$

Laboratory B

Indicator	Measure	Time period	Weight	Benchmark
1. Scientific technical impacts on direct user of R & D outcomes	(0) new/improved products	9/93	80	NA
	(0) improved tests	9/93	20	NA
2. Economic impacts on diirect users of R & D outcomes	(0) actual cost reductions	9/92–9/93	50	NA
	\$100 K actual improvement in a technique for materials handling process	9/92–9/93	50	NA

By applying equation (3):

$IV_1 = 0$ (weight assigned = 0.8)

$IV_2 = 50$ (weight assigned = 0.2)

By applying equation (4):

$$\text{LOI(B)}_2 = 10.0 \quad [\text{Benchmark} = \text{NA}]$$

(3) KOI for Laboratory A

Weights assigned:

LOI(A)_1 (immediate outputs) = 0.3

LOI(A)_2 (intermediate outputs) = 0.7

From equation (5):

$\text{KOI}_{(A)} = (10.04) \times 0.3 + (8.42 \times 0.7) = \mathbf{8.90}$.

(4) KOI for Laboratory B

Weights assigned:

LOI(B)_1 (immediate outputs) = 0.8

LOI(B)_2 (intermediate outputs) = 0.2

From equation (5) above:

$\text{KOI}_{(B)} = (28.02) \times 0.8 + (10.0) \times 0.2 = \mathbf{24.41}$.

(5) Analysis

The data in the tables and the weights assigned to indicators and output stages show that Lab A is more practically oriented than Lab B, which is more concerned with scientific R & D and much less with its impacts downstream as recipients of its outputs. In the stage of immediate outputs Lab B scores higher than Lab A, and even higher than its benchmark. In the stage of intermediate outputs, Lab B scored somewhat higher than Lab A due to one case (in 1992/3) in which one idea from a scientist in Lab B produced a major improvement in a technique for materials handling. This was a 'one shot' occurrence. The data are summarized in Table A1.

Table A1. Summary of illustrative indexes of leading outputs

	Benchmark
Laboratory A	
$\text{LOI(A)}_1 = 10.04$	11.48
$\text{LOI(A)}_2 = 8.42$	NA
KOI(1) = 8.90	
Laboratory B	
$\text{LOI(B)}_1 = 28.02$	21.75
$\text{LOI(B)}_2 = 10.00$	NA
LOI(B) = 24.41	

The example thus generated the following findings:[14]

(a) The KOI for Lab B is much higher than Lab A. However, a multi-year scoring may indicate that the difference between the labs may be smaller or reversed.

(b) Due to the nature of its R & D, Lab B has had a higher performance in the stage of immediate outputs. However, without the one case of economic impact, Lab B would have scored much lower than Lab A in the stage of intermediate outputs. The aberration of 1992/3 for Lab B may be corrected in multiple measures over several years.

(c) Lab A scored at or above its benchmarks in all core indicators, and below benchmarks in the organization-specific indicators. This may be due to the bias that exists in the federal laboratories network of assigning high value to scientific expertise and outcomes at the expense of co-operation with industry.[15]

(d) The KOI for Lab B is almost triple that of Lab A, due perhaps to the factors listed above. However, although a single year's assessment by LOI_i may have intrinsic biases, Lab B has definitely a higher performance rate—as measured in this model. Factors such as preference in the federal laboratories for scientific performance are *external* to the biases of the model itself and would thus influence *any* model of performance evaluation of R & D in these laboratories.

Notes

(1) This is an illustration of an exploratory study of two federal R & D laboratories in the US. The laboratories kindly allowed the author to interview 31 of their employees (12 managers and 19 bench-scientists) for the generation of indicators and measures. The laboratories are of different sizes, and also differ in their governmental affiliation, areas of science/technology, and geographical location. Both laboratories practice an array of fundamental as well as applied research. Preliminary results from this study of the two laboratories were originally reported in [14].

(2) This reflects outputs generated between September 1992 and September 1993.

(3) Weights were assigned by the respondents in each laboratory to reflect the importance of the indicator. Ratings are then normalized to sum to 1.0.

(4) Standards or benchmarks for each measure and indicator were determined by the sample of respondents in one of the following ways: (1) historical data of performance in recent past; (2) perceptions of senior management of what constitutes superior performance (based on personal expectations, experience, performance of top competitors; and/or expectations of the sponsoring federal agency and the language of the mission of the laboratory).

(5) This is a measure of proximity to a pool of scientists and engineers and the research and instructional facilities of an institution of higher learning. Geographic proximity allows for better interface with the scientific community as well as opportunities for training and updating for the laboratory's staff.

(6) Due to the fact that the higher the distance, the lower the attractiveness, this measure was computed on a 3 point rate, in which 1 = low proximity and 3 = high proximity—thus obtaining a score of a positive number to facilitate calculations. A colleague has suggested that all measures should be on a Likert-type scale. Such a procedure would, however, add another cumbersome rating step with additional potential biases.

(7) This measure was not suggested in laboratory A. Different measures may be applied to different organizations as they reflect the organization's choice of measures in view of the nature of its R & D. Interorganizational comparisons will not be biased because of normalization effects of the different weights assigned to the measures *and* to the indicators.

(8) The values for the indicators, measures, and weights have been provided by the sample of 31 employees of the laboratories. Although these values represent the perceptions of the laboratories' personnel of outputs and processes in the organizations *receiving* their outputs and *transforming* them, we utilized these values without major bias. Several measures are known to the laboratory because the receiving organizations provide such reports and the laboratory needs to ascertain its success in technology transfer. Also, to make the example applicable across other types of R & D organizations, the measure of CRADAs (Cooperative R & D Agreements) was not included. This measure is usually used by laboratories for immediate outputs. Clearly a much more accurate mode would be interviews in the organizations producing the intermediate outputs. The exploratory nature of the study, however, did not allow for such additional interviews.

(9) The 7 R & D managers and 9 bench scientists interviewed in laboratory A selected one commercial company (for illustrative purposes) to which their outputs had been transferred. All values for indicators and their weights are based on the interviewees' knowledge and perceptions of this company's process.

(10) This time period is the point in time when measurements are taken.

(11) Benchmark data were obtained from the sample of laboratory managers and scientists who provided their perception of what the company would consider the company's outcomes benefitting from the laboratory's transfer of R & D outputs to the company.

(12) Not available. Laboratory personnel did not know or would not advance this information.

(13) For intermediate outputs there was no distinction between the core and organization-specific indicators.

(14) Since this is an exploratory study, additional findings may be generated with a more detailed implementation of the model in a larger and more comprehensive study.

(15) It is worthwhile to note that many federal laboratories have inherent in their mission some of the downstream output indicators such as: contribution to national health and national defense.

APPENDIX B

Illustrative Development of Indexes of Cost-performance of R & D for Immediate and Intermediate Outputs

(1) Immediate outputs

Laboratory A

From equation (6):

Index of Cost Performance (ICP)

$$= \frac{[10.04]}{\text{Cost of R \& D}}.$$

Cost of R & D was computed as the total expenditures for R & D over a period of the past 2 years. This measure was selected because it is the average time for completion of an R & D project in the laboratory. Thus, expenditures were computed for FY 1991 and 1992, so that:

$$\text{ICP} = \frac{[10.04]}{\$5.8\ \text{(million)}} = [1.73].$$

The ICP per cost of S & E was computed so that

$$\text{ICPSE} = \frac{[10.04]}{\$88\ \text{(averaged over the 2 years)}} = [0.11].$$

Laboratory B

From equation (6):

$$\text{ICP} = \frac{28.02}{\text{cost of R \& D}}.$$

Cost of R & D was computed on the total expenditures for R & D in the laboratory over the past 4 years. This is the average time for completion of a project in the laboratory, so that:

$$\text{ICP} = \frac{28.02}{\$42\text{(million)}} = 6.67.$$

The ICP per cost of Scientist and Engineer was computed on the average cost per S & E over the past 4 years, so that:

$$\text{ICPSE} = \frac{28.02}{\$96\ \text{(average over 4 years)}} = 0.29.$$

(2) Intermediate outputs

The intermediate outputs stage of the ICP was not computed due to lack of data on expenditures in the companies to which the outputs from the laboratories had been transferred.

(3) Analysis

Lab B is much smaller than Lab A with a more expensive staff of scientists and engineers (three times more doctorates than A). Lab B seems to be more productive (per dollar expended and per cost of S & E) than Lab A. This finding holds over for ICPSE for Lab B, where the cost per scientist and engineer is higher than in Lab A. Notwithstanding the biases discussed in Appendix A, Lab B seems to be more cost effective than Lab A.

APPENDIX C

Illustrative Link of Downstream Output Indicators to Individual Research Effort on the R & D/Innovation Flow

The rationale for linking downstream output indicators (pre-ultimate and ultimate) to individual R & D effort is described in papers by Rubenstein and Geisler [35, 36], and Geisler [14].

In Appendix C there is an illustration of one R & D project in Lab A in the example in Appendixes A and B. Following Rubenstein and Geisler [35], a flow diagram was constructed for the project code name Alpha. The project entailed development of a substance from organic sources with capabilities of absorption of liquids beyond those of existing synthetic materials.

The flow diagram in Fig. C1 shows the output indicators for each stage, and the barriers and facilitators for each diffusion and transformation step in the process. The data are based on interviews with laboratory personnel involved and familiar with project Alpha. Data on downstream transformation of the outputs are based on perceptions, recollections and 'best guesses' of respondents in the laboratory. These data may be confirmed and reassessed through interviews in the respective organizations downstream the R & D/innovation flow.

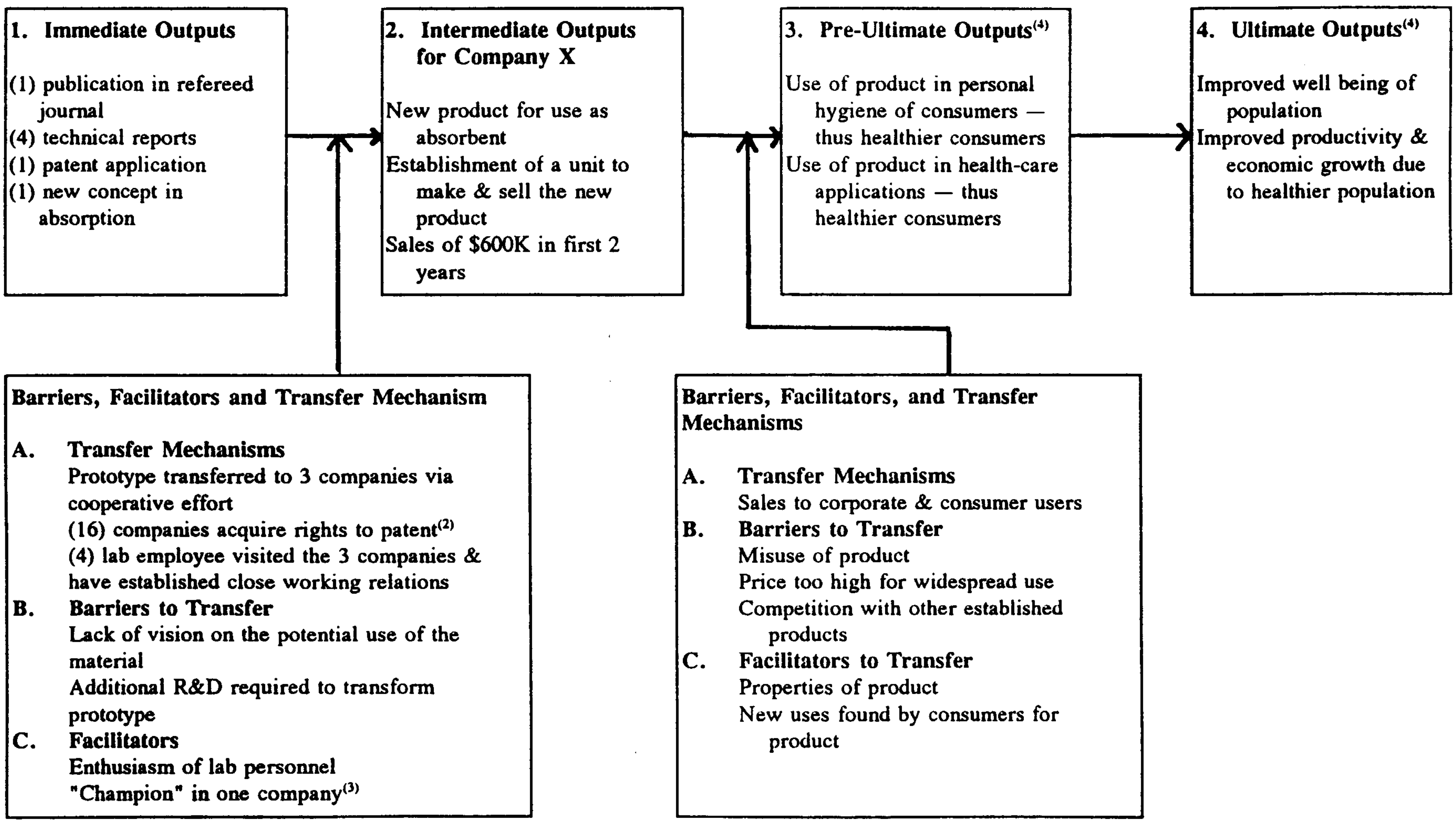

Fig. C1. Flow diagram for Project Alpha.[1] [1]For proprietary reasons and desire for anonymity, all details of the project and the product have been disguised. It is possible that a product such as the one described here may exist in a federal laboratory. [2]Rights to patents on non-secret outcomes from federal laboratories in the US may be purchased by multi-users for a nominal fee. In some instances laboratories have entered into sole user agreements with private companies on outcomes from jointly funded R & D. [3]In this case, one company had a 'champion' in one of the R & D managers who saw the potential uses and had allocated corporate R & D funds for additional work on the prototype—leading to a new product, new unit, and sales of $600 K in the first two years. Although for a chemical company such as X these sale figures are puny, it is nevertheless a case-study of a successful product emerging from divisional R & D and attributed to output from a federal R & D laboratory. [4]Due to the inability to interview respondents in the company and other downstream entities (in a 'forward' process akin to the hindsight approach [9]), these outputs lack financial or other quantitative measures for their indicators.

ACKNOWLEDGEMENTS

This research was supported by the National Science Foundation (Grant No. SBR-9401432) and by the College of Business and Economics at the University of Wisconsin-Whitewater. Special thanks are due to the anonymous respondents in the research sites, and to three anonymous referees for their insightful suggestions. Also, special thanks are due to James Felli for his insights and assistance in data analysis.

REFERENCES

1. Baltagi B and Griffin J (1988) A general index of technical change. *J. polit. Econ.* **96,** 20–41.
2. Bobe B (1991) Trends in the use of R & D output indicators in EC program evaluation. *Scientometrics* **21,** 263–282.
3. Bozeman B and Melkers J (Eds) (1993) *Assessing R & D Impacts: Methods and Practice.* Kluwer Academic, Norwell, Mass.
4. Brown W and Gobeli D (1992) Observations on the measurement of R & D productivity: a case study. *IEEE Trans. Engng Mgmt* **39,** 325–331.
5. Callon M, Law J and Rip A (Eds) (1986) *Mapping the Dynamics of Science and Technology.* Macmillan, London.
6. Capron J (1992) At state-of-the-art of quantitative methods for the assessment of R & D programs. In *Management of Technology* (Edited by Khalil T and Bayaktar B), pp. 1195–1204. Institute of Industrial Engineers.
7. Chubin D (1987) Research evaluation and the generation of big science policy. *Knowledge* **9,** 254–277.
8. Cordero R (1990) The measurement of innovation performance in the firm: an overview. *Res. Policy* **19,** 185–192.
9. DoD, *Project Hindsight.* Office of the Director of Defense Research and Engineering, Washington, D. C., DTIC No. AD495905, 10169.
10. Dryden J (1992) Quantifying technological advance: S&T indicators at the OECD—challenge for the 1990s. *Sci. Publ. Policy* **19,** No. 5, 281–290.
11. Evered D and Harnett S (Eds) (1989) *The Evaluation of Scientific Research.* Wiley, New York.
12. Fernelius W and Waldo W (1980) Role of basic research in industrial innovation. *Res. Mgmt* **23,** No. 4, 36–40.
13. Geisler E (1995) The cost of research. In press.
14. Geisler E (1995) Key output indicators in performance evaluation of R & D organizations. *Technol. Forecast. Social Change* **47,** In press.
15. Gold B (1989) Some key problems in evaluating R & D performance. *J. Engng Technol. Mgmt* **6,** 59–70.
16. Griliches Z (Ed.) (1984) *R & D, Patents and Productivity.* University of Chicago Press.
17. Holbrook JA (1992) Basic indicators of scientific and technological performance. *Sci. Publ. Policy* **19,** 267–273.
18. Irvine J, Martin B, Abraham J and Peacock T (1987) Assessing basic research: reappraisal and update of an evaluation of four radio astronomy observatories. *Res. Policy* **16,** 213–227.
19. Jensen E (1987) Research expenditures and the discovery of new drugs. *J. Indust. Econ.* **36,** 83–95.
20. Lederman L (1984) The value of fundamental science. *Scient. Am.* **251,** No. 5, 34–41.
21. McGrath M and Romeri M (1994) The R & D effectiveness index: a metric for product development performance. *J. Prod. Innov. Mgmt* **11,** 213–220.
22. Mansfield E, Romeo A and Switzer L (1983) R & D price indexes and real R & D expenditures in the United States. *Res. Policy* **12,** 105–112.
23. Mansfield E (1991) Academic research and industrial innovation. *Res. Policy* **20,** 1–12.
24. Mansfield E (1992) Academic research and industrial innovation. A further note. *Res. Policy* **21,** 295–296.
25. Martin B, Irvine J, Narin F and Stevens K (1990) Recent trends in the output and impact of British science. *Sci. Publ. Policy* **17,** 14–26.
26. Mechlin G and Berg D (1980) Evaluating research—ROI is not enough. *Harvard Bus. Rev.* Sept–Oct. 93–100.
27. Morbey G (1988) R & D: its relationship to company performance. *J. Prod. Innov. Mgmt* **5,** 191–200.
28. Mullins N (1987) Evaluating research programs: measurement and data sources. *Sci. Publ. Policy* **14,** No. 2, 56–68.
29. Narin F (1987) Bibliometric techniques in the evaluation of research programs. *Sci. Publ. Policy* **14,** No. 2, 99–106.
30. Nederhof A and Van Roan A (1993) A bibliometric analysis of six economics research groups: a comparison with peer review. *Res. Policy* **22,** 353–368.
31. Nelson R (1959) The simple economics of basic scientific research. *J. Polit. Econ.* **67,** 297–306.
32. Palda K (1986) Technological intensity: concept and measurement. *Res. Policy* **15,** 187–198.
33. Park J and Chong J (1991) A model to assess the value of an intermediate R & D result. *IEEE Trans. Engng Mgmt* **38,** 157–163.
34. Robb W (1991) How good is our research? *Res. Technol. Mgmt* **34,** No. 2, 16–21.
35. Rubenstein AH and Geisler E (1991) Evaluating the outputs and impacts of R & D/innovation. *Int. J. Technol. Mgmt* **6,** 181–204.
36. Rubenstein AH and Geisler E (1988) The use of indicators and measures of the R & D process in evaluating science and technology programs. In *Government Innovation Policy* (Edited by Roessner J), pp. 185–204. St. Martins Press.
37. Schmied H (1987) About the quantification of the economic impact of public investment into scientific research. *Int. J. Technol. Mgmt* **2,** 711–730.
38. Sen BI and Shailendra K (1992) Evaluation of recent scientific research output by a bibliometric method. *Scientometrics* **23,** 31–46.
39. Tijssen RJ (1992) A quantitative assessment of interdisciplinary structures in science and technology: co-classification analysis of energy research. *Res. Policy* **21,** 27–44.
40. Trajtenberg M (1990) A penny for your quotes: patent citations and the value of innovations. *Rand J. Econ.* **21,** 172–187.
41. U.S. National Science Board (1993) *Science and Engineering Indicators, 1993,* 11th Edition. National Science Foundation, Washington, D. C.
42. Utterback JM and Abernathy W (1975) A dynamic model of product and process innovation. *Omega* **3,** 639–656.
43. Van Raan AF (1988) (Ed.) *Handbook of Quantitative Studies of Science and Technology.* North-Holland Elsevier, Amsterdam.
44. Von Winterfeldt D and Edwards W (1986) *Decision Analysis and Behavioral Research.* Cambridge University Press, New York.
45. Zenger TR (1994) Explaining organizational diseconomies of scale in R & D: agency problems and the allocation of engineering talent, ideas, and effort by firm size. *Mgmt Sci.* **40,** 708–729.

ADDRESS FOR CORRESPONDENCE: *Professor E Geisler, Department of Management, College of Business and Economics, University of Wisconsin, Whitewater, WI 53190-1797, USA.*