MITSloan

The International Center for Research on the Management of Technology

Metrics to Evaluate R,D&E

John R. Hauser
Florian Zettelmeyer*

October 1996 WP # 156-96

Sloan WP # 3934

*Assistant Professor of Marketing
University of Rochester
Simon Graduate School of Business Administration
Rochester, New York

Sloan School of Management
Massachusetts Institute of Technology
38 Memorial Drive, E56-390
Cambridge, MA 02139-4307

Metrics to Evaluate R,D&E

by

John R. Hauser

and

Florian Zettelmeyer

October 15, 1996

This paper has been prepared for *Research Technology Management* as a research briefing. The material in this paper summarizes research that appears in "Evaluating and Managing the Tiers of R&D," "Metrics to Value R&D Groups, Phase I: Qualitative Interviews," and "Metrics to Value R&D: An Annotated Bibliography," all M.I.T. working papers by the authors. These papers are available from the International Center for Research on the Management of Technology (ICRMOT) at M.I.T.'s Sloan School of Management, Cambridge, MA 02142. To review and/or order these and other ICRMOT working papers, visit our web page at http://web.mit.edu/icrmot/www/.

John R. Hauser is the Kirin Professor of Marketing, Massachusetts Institute of Technology, Sloan School of Management, 38 Memorial Drive, E56-314, Cambridge, MA 02142, (617) 253-2929, (617) 258-7597 fax, jhauser@sloan.mit.edu. Florian Zettelmeyer is an Assistant Professor of Marketing at the University of Rochester, Simon Graduate School of Business Administration, Rochester, NY 14627, (716) 271-5522, florian@mail.ssb.rochester.edu.

Metrics to Evaluate R,D&E

OVERVIEW. Corporate downsizing has brought increasing pressure on R,D&E managers to develop metrics with which to value the return on R,D&E investments. But "you are what you measure." Metrics affect research decisions, research efforts, and the researchers themselves. The authors summarize the insights they have obtained with respect to R,D&E metrics. These insights are based on a review of the literature, interviews with 43 representative CEOs, CTOs, and researchers at ten research-intensive organizations, and formal mathematical analyses. The best metrics depend upon the goals of the R,D&E activity as they vary from applied projects to competency-building programs to basic research explorations.

For applied projects, market outcome metrics (sales, customer satisfaction, margins, profit) are relevant if they are adjusted via corporate subsidies to account for short-termism, risk aversion, scope, and options thinking. The magnitude of the subsidy should vary by project according to a well-defined formula.

For R,D&E programs which match or create core technological competence, outcome metrics must be moderated with "effort" metrics. Too large a weight on market outcomes leads to false rejection of promising programs. The large weight encourages the selection of lesser value programs that provide short-term, certain results that are concentrated in a few business units. This, in turn, leads a firm to use up its "research stock." Instead, to align R,D&E with the goals of the firm, the metric system should balance market outcome metrics with metrics that attempt to measure research effort more directly. Such metrics include many traditional indicators.

For long-term research explorations, the right metrics encourage a breadth of ideas, For example, many firms seek to identify their "best people" by rewarding them for successful completion of research explorations. However, metrics implied by this practice lead directly to "not-invented-here" attitudes and result in research empires that are larger than necessary, but lead to fewer total ideas. Alternatively, by using metrics that encourage "research tourism," the firm can take advantage of the potential for research spillovers and be more profitable.

In a recent issue of *Research Technology Management*, Arthur Chester (1), Senior Vice President for Research and Technology at GM Hughes Laboratories, stated that: "measuring and enhancing R&D effectiveness ... has gained the status of survival tactics for the R&D community." This sentiment was echoed as an important policy issue in Japan (2) and Europe (3). Research, Development, and Engineering (R,D&E) metrics are important for at least three reasons. First, such metrics document the value of R,D&E and are used to justify investments in this fundamental, long-run, and risky venture. Second, good metrics enable Chief Executive Officers (CEOs) and Chief Technical Officers (CTOs) to evaluate people, objectives, programs, and projects in order to allocate resources effectively. Third, metrics affect behavior. When scientists, engineers, managers, and other R,D&E employees are evaluated on specific metrics they make decisions, take actions, and otherwise alter their behavior in order to improve the metrics. The right metrics align employees' goals with those of the corporation. The wrong metrics are counterproductive and lead to narrow, short-term, and risk avoiding decisions and actions.

The International Center for Research on the Management of Technology (ICRMOT) at M.I.T.'s Sloan School of Management is funding an ongoing scientific study of R,D&E metrics in order to understand and improve their use in industry. This research briefing describes what we have learned to date. We began with in-depth interviews with 43 representative CEOs, CTOs, and researchers at ten research-intensive organizations (4) including Chevron Petroleum Technology, Hoechst Celanese ATG, AT&T Bell Laboratories, Bosch GmbH, Schlumberger Measurement & Systems, Electricite de France, Cable & Wireless plc, Polaroid Corporation, US Army Missile RDEC and Army Research Laboratory, and Varian Vacuum Products. All interviews were conducted in the native languages of the managers. We continued with a comprehensive review of the published literature (5). We then attempted to abstract and generalize the insights we obtained from these exercises. This resulted in a scientific theory to guide the selection of R,D&E metrics (6).

We attempt here to summarize the basic intuition resulting from our research. We hope to highlight the long-term implications of current trends and suggest improvements to current practice. If nothing else, we seek to encourage debate on issues that are fundamental to managing R,D&E.

The Tier Metaphor

R,D&E is a diverse activity. Some applied projects attempt to solve short-term problems faced by a single business unit. There may be little uncertainty as to the outcome of these projects because they are well within the current capabilities of the R,D&E organization. At the other extreme, some research explorations seek to build a basic competency in an area of science that is likely to be important to the corporation in the future. These explorations often have a much longer perspective, apply to many business units, and entail considerable risk. Other programs fall somewhere between these extremes. Naturally, the mix of projects, programs, and explorations varies by organization. For example, central R&D laboratories might have a greater percentage of basic research programs than R,D&E operations within a business unit. Indeed, many individual scientists and engineers have a mix of activities within their own portfolios. However, our observations suggest that every organization faces the challenge of integrating these diverse activities.

Best-practice organizations recognize variation and select metrics accordingly. By selecting the right metric for each activity the firm encourages the right decisions and actions by scientists, engineers, and managers. If a firm applies the same metrics throughout the R,D&E process, they do not get the most out of their technological efforts. Many of the mistakes that we observed in the field occurred when technology managers attempted to apply the same metrics throughout the process.

To understand better how metrics vary, we introduce a tier metaphor. This metaphor enables us to categorize a diverse continuum of projects, programs, and explorations and focus on key characteristics. We define "tier 1" as basic research which attempts to understand basic science and technology. Tier 1 explorations may have applicability to many business units. Indeed, they may spawn new business units. We define "tier 2" as those activities which select and develop programs to match or create the core technological competence of the organization. "Tier 3" is defined as specific projects focused on the more immediate needs of the customer, the business unit, and/or the corporation. Tier 3 is often accomplished with funding by both business units and the corporation. For clarity we adopt the terminology of (7) and use the words "objectives" and/or "explorations" for tier 1 activities, the word "programs" for tier 2 activities, and the word "projects" for tier 3 activities. We note, however, that these terms are often used interchangeably in the literature and in practice.

While the word "tier" was used at only a few of the organizations we visited, most firms had a concept that the management of technology varied depending on the stage of the process. See also (8), (9), (10), and (11). For example, the U.S. Army classifies their research as 6.1, 6.2, and 6.3 (and beyond) which corresponds to our metaphor of tiers 1, 2, and 3. Even the shorthand, RD&E seems to separate the stages of innovation.

We now use the tier metaphor to illustrate how the issues of short-termism, risk aversion, option value thinking, scope, portfolio planning, research spillovers, and research tourism affect R,D&E metrics. While these issues are relevant to all R,D&E activities, the implications and foci of these issues are more intense in some tiers than others. We begin with tier 3, that is, projects that address shorter-term issues with one or more well-defined customers.

Should R,D&E Projects be Entirely Customer Driven?

Many managers, consultants, and researchers have argued that, to succeed, R,D&E should be more customer-driven. This viewpoint was reinforced in our interviews. "R,D&E has to be developed in the marketplace." "Technical assessment is 'What does it do for the customer?'" In many instances R,D&E managers maintained their budgets by "selling" projects to internal customers such as the business units.

There is no doubt that the customer is important. In order for the firm to have a good bottom-line (profits) it must have a good top-line (revenue). Good top-line performance means products and services that are designed to fulfill customer needs and satisfy customers (13). R,D&E provides the means with which the firm achieves good top-line performance.

However, our interviewees recognized a downside to a pure customer focus. Instead, many subsidized R,D&E projects with central funds. Business units were asked to pay only a fraction of the cost of an R&D project. One CTO stated that the business units were better able to judge an R,D&E project if they did not have to pay the full cost. One business unit manager told us about "tin cupping" where she would go around to other business unit managers to ask for contributions to a research project as if she were a beggar with a tin cup. These firms have structured themselves so that the customer is an arbiter, but not the only arbiter, of R,D&E funding.

In addition, there is scientific evidence supporting a perspective that all projects should not be entirely customer-driven. Mansfield (12) found that, holding total R&D expenditure constant, an organization's innovative output was driven by the percent allocated to basic research. Cooper and Kleinschmidt (14) found that adequate resources in R,D&E was a key driver separating successful firms from unsuccessful firms. And, Bean (15) suggests that a greater percentage of research activities driven by R&D implies more growth.

In our interviews and analyses we found that both viewpoints have merit. Good project decisions balance customer-driven and research-driven foci. We found that, with the proper central subsidies and "options thinking," business units could select the R,D&E projects that were in the best interests of the firm. However, without well-designed subsidies there was a bias toward short-term, narrow projects with predictable outcomes. We address each of these issues in turn.

Short-termism

For an R,D&E manager the costs of R,D&E projects are easy to observe and occur "today," but the benefits of R,D&E projects are realized many years in the future and may not be attributed to the project. Furthermore, because good scientists, engineers, and managers are mobile, they may leave their jobs, or even the firm, before market outcomes such as customer satisfaction, sales, or profit can be observed. Even if they stay in the same job, they may not get credit for the benefits. However, scientists, engineers, and managers face mortgages, tuition bills, and other expenses that occur now and are more predictable than long-term customer-based measures. It is rational for these employees to be more short-term oriented than the firm.

The impacts of such individual rationality can be dramatic. For example, a ten-year project might be valued by the firm at $100 million given its discount rate but only $86 million by a business unit manager with a discount rate only 1% higher.[1] Thus, the project has a value to the business unit manager that is only 86% of the value to the firm. We call this a short-termism ratio of $\gamma=0.86$.

Furthermore, the short-termism ratio varies by project. The impact of short-termism is more dramatic for projects with long-term payback -- the short-termism ratio might be 0.86 for a project with a rapid payback but 0.50 for a project with a payback spread over many years. If the firm does not use central subsidies to adjust for this effect, then there will be strong business unit pressure to fund only the short-term projects. The same phenomenon applies to any rewards, incentives, or evaluations of individual scientists, engineers, or managers in R,D&E. Because R,D&E employees discount the future more than the firm, unadjusted customer-based measures will cause them to favor short-term oriented projects.

Risk Aversion

Some projects are more risky than others. A large firm can diversify this risk across many projects and stockholders can diversify risk across firms. But individual business-unit managers can not diversify risk as easily. If they are risk averse, and most managers are, they will undervalue risky

[1]For this example take a time stream of profits (in $millions) consisting of -9, -12, -20, -8, 0, 5, 14, 20, 28, 35, 38, 40, 41, 42, and 43. Use discount rates of 7% for the firm and 8% for the business unit.

projects. For most situations we can approximate the effect of risk aversion with a ratio similar to the short-termism ratio. We call this ratio R (for risk aversion).[2] See (6) for details. Without central subsidies to counteract this risk factor, business units tend to favor projects with predictable paybacks even though the projects provide less value to the firm.

Scope

Even if only one business unit funds an applied project, many business units might benefit. For example, Mechlin and Berg (16) illustrate scope by discussing research at Westinghouse on water flows through porous geological formations. This research was done for the uranium mining division, but had substantial additional benefits for heat-flow analyses for high-temperature turbines and below-ground heat pumps, and for the evaluation of environmental impacts (real estate division). Such scope was highlighted many times by our interviewees. The firm, but not the business unit, realizes the benefits of scope.

Thus, we add one more ratio -- a concentration ratio (α) -- to reflect the percent of total benefits that accrue to the business unit for which the R,D&E project was completed. (For example, if approximately 40% of total benefits accrue to the business unit providing the funding, then α = 40%.) If the firm does not adjust for such concentration, R,D&E's customers will favor focused projects over those that have wide applicability.

Options Thinking

The last concept highlighted by our interviewees was options thinking. The idea is simple. Investing in an R&D project is like buying a financial option to make further investments. If, as the result of initial investigations, further investment is justified, the firm will invest further in the project. (In theory) if further investment is not justified, the firm will abort the project. This means that the value of the project should reflect these investment contingencies -- the option value is higher than

---

[2]Technically, R is the ratio of the certainty equivalent of the income stream to the expected value of the income stream. The certainty equivalent is the amount of guaranteed income the business unit would accept in place of the uncertain income stream. For example, if the income stream were relatively uncertain, the business unit manager might accept a guaranteed income stream that was only 90% as large. In this case, R=90%.

that which would be calculated if all future investments were locked in. Options thinking implies that outcomes uncertainty provides an option value. For more discussion, see Faulkner (17).

Options thinking is critical to project evaluation, but it applies equally to the firm and to the business unit. Thus, it does not affect the calculation of central subsidies beyond that already captured by the short-termism and risk aversion indices.

Variation in Subsidies

We found that the magnitude of the subsidy varied by firm and, in some cases, by project. The variation was important because the characteristics of payback-length, risk, and concentration varied. Setting the best subsidy for a project was recognized as a critical management challenge. Interestingly, if one interprets "tin-cupping" as an auction in which business units "bid" for R,D&E projects, then tin-cupping might be an efficient economic means by which the firm can overcome the short-termism, risk aversion, and narrow foci of business unit managers.

Summary

We found that firms subsidize R,D&E projects in order to align internal customer decisions with those of the firm. When used properly, these subsidies adjust for variations in short-termism, risk aversion, and scope. The optimal subsidy (S) is given by the equation: $S = \gamma R \alpha$, where γ is the short-termism ratio, R is the risk aversion index, and α is the concentration index.

More importantly, from a research policy perspective, firms should retain central subsidies of research projects. Such subsidies maintain a long-term, wide-scope, balanced mix of R,D&E projects. A more subtle message is that S varies by project. Firms which use a single subsidy ratio (and many do) gain in ease of implementation, but they are not achieving the efficiency that is possible with a more flexible process.

Metrics for Selecting Technology to Match or Create Core Technological Competence

We now consider R,D&E activities that attempt to select technology to match or create core technological competence (tier 2). This is an important function of R,D&E. On one hand, these programs are based on the firm's strategic plans and technological capabilities. On the other hand, these programs determine future core technological competence. Our interviewees suggested that it is a critical challenge to decide how heavily customer input should be weighed in the selection of tier 2 programs. As one of our interviewees said: "The customer knows the direction but lacks the expertise; researchers have the expertise, but lack the direction."

Some firms are attempting to make these decisions customer-driven by using metrics that measure "outcomes," that is, sales, satisfaction, or incremental profit. For example, one popular metric measures R,D&E effectiveness by comparing the profit due to new products to the amount spent on R,D&E (18). Many of our interviewees and our analyses suggest that such metrics, when used alone, increase profits in the short-term, but sacrifice the future.

To illustrate this phenomenon, we consider a slightly stylized representation of the process by which R,D&E decides to invest in science and technology in order match or create core technological competence. This stylized representation is based on suggestions by our interviewees that the most critical decision in tier 2 was to <u>select</u> the right programs. Motivating the right amount of scientific, engineering, and process effort was important, but not as important as program selection. These statements suggest the following process.

<u>Step 1</u>. R,D&E chooses one or more programs to develop science and technological capabilities to fulfill (or anticipate) customer needs. Such programs develop resources for competitive advantage.

<u>Step 2</u>. R,D&E undertakes initial research to evaluate the program(s). This research determines the potential contribution if the program is successful.

<u>Step 3</u>. R,D&E invests scientific, engineering, and process <u>effort</u> to refine the program into one or more applied projects. This effort matches R,D&E capability to the needs of the business units.

The order of the steps is important. Program choice is made before the value of the research is known and before the bulk of the research activity is undertaken. Thus, R,D&E faces considerable uncertainty in the outcomes of program choice. If the firm uses metrics which encourage R,D&E to base program choice on anticipated market outcomes, then the impacts of risk aversion and short-termism are enormous. The effects are so large that the firm can not overcome these effects with program subsidies alone. (This is one way in which programs differ from projects.)

Market-based Outcome Metrics vs. Choosing the Best Program

Risk aversion and short-termism lead to false rejection -- some programs are rejected that are of value to the firm, and false selection -- short-term, certain programs are favored relative to long-term, risky programs that have a larger expected value to the firm. When we mapped the regions for false rejection and false selection we were surprised at the strength of these effects for program selection. We found that the only way to avoid large regions of false rejection and false selection was to place a low weight on market outcome metrics, that is, low relative to other metrics.

Our interviews and analyses are at odds with recent authors (e.g., 18) who advocate a simple comparison of market outcomes to research costs. Such metrics could lead to decisions and actions that use up "research stock" by favoring short-term, certain projects. If the firm does not recognize that programs to match or create core technological competence differ from applied projects, then a heavy emphasis on outcome metrics will lead to under-investment in new technologies and science. This, in turn, could lead to long-term ruin.

Balancing Effort and Outcomes After Program Choice

Despite their problems for these R,D&E activities, we can not reject market-based outcome metrics entirely. Not only must step 1 have some input from the market, albeit small relative to other metrics, but outcome metrics are critical to step 3. In step 3, the firm's goal is to motivate scientists, engineers, and managers to allocate the right amount of scientific, engineering, and process effort after the program has begun. R,D&E must incur costs and the people involved must be motivated to work on the projects that are best for the firm. Because these costs are real, individual decisions will only be aligned with the goals of the firm if the rewards are aligned with the rewards to firm. To see this another way, consider that costs and individual efforts are easy for scientists, engineers, and managers to observe. This means that such costs are given a high implicit weight in any decision. If the firm wants to balance outcomes and costs, it must provide metrics to assure that R,D&E gives the proper weight to outcomes. These arguments imply a larger weight on market-based outcome metrics.[3]

We now face a management dilemma. To encourage the right program choice, the firm wants a small weight on market-based outcomes metrics. To motivate the right allocation of effort after program choice, the firm wants a large weight on market-based outcomes metrics.

In theory, a firm can overcome this dilemma if it can find metrics that measure today's RD&E efforts without exposing RD&E to the risks inherent in long-term market outcomes. When such metrics are available, the firm can evaluate R,D&E on the effort metrics (and thus align potential outcomes with costs). Of course this means that the effort metrics must correlate with expected, long-term market outcomes. By placing a larger weight on the effort indicators for tier 2 programs and a smaller weight on market outcomes for tier 2 programs, the firm attempts to balance the motivations for the right decisions (small weight on market outcomes) and the right effort (larger weight on effort metrics). With the right balance, at least in theory, scientists, engineers, and managers will, acting in their own best interests, select the programs that are best for the firm and allocate the right amount effort to complete those programs.

Unfortunately, we found few ideal "effort" metrics. The firms we interviewed attempted to use metrics such as publications, citations, patents, citations to patents, peer review and other measures as indicators of scientific and engineering effort. However, each metric had potential problems because each metric, taken alone, could be "gamed" by the scientists, engineers, and managers. No one metric captured all relevant efforts. To overcome these deficiencies, most firms used a combination of metrics.

Summary

The recent trend toward a heavy reliance on customer-driven outcome metrics (sales, satisfaction, profit) is counter-productive because it sacrifices long-term benefits in the development of core technological competence. This effect is most pronounced in the choice of science and technology programs. However, some metrics are needed to motivate the right amount of scientific,

[3]It is possible to derive these implications mathematically with a set of methods know as "agency theory." The basic idea is to set up equations for how the "agents" will react to the metric system and then adjust the metrics until the agents, acting in their own best interests, choose those actions and make those decisions that are in the best interests of the firm. For details see (6). These equations balance the effects of outcome metrics, cost metrics, risk, and short-termism.

engineering, and process effort. The best metric system uses a combination of outcome and "effort" metrics. The longer term and more risky the research, the lower the weight should be on market-based outcomes. To the extent that they measure effort, we should not reject traditional metrics, such as publications, citations, patents, citations to patents, and peer review. However, they should be used in combination to minimize "gaming" effects.

Research Tourism vs. Not Invented Here

We turn now to the basic research explorations which provide the scientific and technological knowledge upon which tier 2 programs are based. We call these activities tier 1 explorations.

Basic research explorations are the most difficult to measure. Not only is the outcome of scientific investigations unknown, but specific business implications are difficult to predict. Furthermore, researchers often have a better idea (than management) of which explorations will be in the best interests of the firm's core technological strategy.

We found many issues in tier 1, such as the selection of the best people and the balancing of a high-variance research portfolio. Because these issues are covered in the extant literature, we do not focus on them here. (See [5] for a review.) Instead, we focus on how some metric systems encourage "research tourism" and others encourage "not-invented-here" decisions.

By research tourism our interviewees referred to a common practice among R,D&E employees of visiting other laboratories and universities and of entertaining visitors from other laboratories and universities. Attending conferences and reading the literature might also be considered research tourism. The business purpose is to identify and evaluate outside ideas that have the potential to enhance a firm's internal development. The literature calls these outside ideas "research spillovers." If research spillovers are managed correctly, they can be quite profitable. For example, in an econometric study Jaffe (19) suggests that the indirect effect of research spillovers from competitors is so large that it more than offsets the fact that competitors' R,D&E strengthens competitors.

However, there is a catch. In other to benefit from research spillovers a firm must maintain its expertise in the area. The more a firm invests in an area, say polymers, the more it is able to benefit from outside research activities in that area.

Despite the importance of research spillovers, we found that many firms identify their best people by the internal explorations that those people complete successfully. When researchers are rewarded mostly for internal explorations, they have less incentive to seek outside ideas.

When we analyzed such metric systems we found that the potential for spillovers make it more profitable for the firm to undertake more total (internal and external) basic science explorations than they would in the absence of spillovers. But we found that the most profitable number of internal explorations might actually decrease. In other words, a policy of seeking outside ideas could cause internal research empires to shrink. When research tourism is encouraged and basic science researchers are measured with respect to the outcome of all explorations, whether internal or external, it is possible to align the incentives of the researchers with those of the firm. Researchers will make the decisions that are in the best strategic interests of the firm.

However, we found that if researchers are measured by internal explorations only, then (1) they will adopt a "not-invented-here" attitude and spend little or no time on research tourism, (2) they will work on more internal explorations than is in the firm's best interests, and (3) the net result will be fewer scientific developments. In other words, not-invented-here is the result of the metrics by which research teams are evaluated rather than a generic property of research teams. Poorly designed metrics lead R,D&E to spend excessive resources on internal ideas and to devote too few resources to external explorations. Such metrics lead to research empires that are larger than they need be.

Fortunately, many RD&E organizations are recognizing the need to reward explicitly ideas that come from outside the firm. For example, in March 1996, the General Motors Corporation approved a vision statement that included the phrase "Deploy more highly valued innovations, no matter their source, than any other enterprise." (Underline added.)

Summary

Table 1 summarizes the recommendations that result from M.I.T.'s ongoing research on R,D&E metrics. In that table we use the tier metaphor to emphasize that a variety of metrics are needed to evaluate and manage R,D&E. Metrics that are best for one type of activity might be counter-productive for another type of activity. We close this briefing with a list (Table 2) of the metrics used by our interviewees. This list is categorized with the tier metaphor. Notice that some metrics measure incremental profit, some are surrogates for incremental profit (e.g., customer satisfaction and time to market), and some attempt to measure scientific and engineering effort.

References

1. Chester, Arthur N. (1995), "Measurements and Incentives for Central Research," *Research Technology Management*, (July-Aug), 14-22.

2. Irvine, John (1988), *Evaluating Applied Research: Lessons from Japan*, (London: Pinter Publishers).

3. European Industrial Research Management Association (1995), *Evaluation of R&D Projects*, Working Group Report No. 47.

4. Zettelmeyer, Florian and John R. Hauser (1995), "Metrics to Value R&D Groups, Phase I: Qualitative Interviews," Working Paper, International Center for Research on the Management of Technology, MIT Sloan School, Cambridge, MA 02142 (March).

5. Hauser, John R. (1996), "Metrics to Value R&D: An Annotated Bibliography," Working Paper, International Center for Research on the Management of Technology, MIT Sloan School, Cambridge, MA 02142 (March).

6. Hauser, John R. and Florian Zettelmeyer (1996), "Evaluating and Managing the Tiers of R&D," Working Paper, International Center for Research on the Management of Technology, MIT Sloan School, Cambridge, MA 02142 (March).

7. Steele, Lowell W. (1988), "What We've Learned: Selecting R&D Programs and Objectives," *Research Technology Management*, (March-April), 1-36.

8. Bachman, Paul W. (1972), "The Value of R&D in Relation to Company Profits," *Research Management*, 15, (May), 58-63.

9. Krause, Irv and Liu, John (1993), "Benchmarking R&D Productivity: Research and Development; Case Study," *Planning Review*, 21, 1, (January), 16-21.

10. Pappas, Richard A. and Donald S. Remer (1985), "Measuring R&D Productivity," *Research Management*, (May-June), 15-22.

11. Tipping, James W., Eugene Zeffren, and Alan R. Fusfeld (1995), "Assessing the Value of Your Technology," *Research Technology Management*, 22-39.

12. Hauser, John R., Duncan I. Simester, and Birger Wernerfelt (1994), "Customer Satisfaction Incentives," *Marketing Science*, 13, 4, (Fall), 327-350.

13. Mansfield, Edwin (1980), "Basic Research and Productivity Increase in Manufacturing," *American Economic Review*, (December).

14. Cooper, Robert G. and Elko J. Kleinschmidt (1995), "Benchmarking the Firm's Critical Success Factors in New Product Development," *Journal of Product Innovation Management*, 12, 374-391.3.

15. Bean, Alden S. (1995), "Why Some R&D Organizations are More Productive than Others," *Research Technology Management*, (Jan-Feb), 25-29.

16. Mechlin, George F. and Daniel Berg (1980), "Evaluating Research -- ROI is Not Enough," *Harvard Business Review*, 59, (Sept-Oct), 93-99.

17. Faulkner, Terrence W. (1996), "Applying 'Options Thinking' to R&D Valuation," *Research Technology Management*, (May-June), 50-56.

18. McGrath Michael E. and Michael N. Romeri (1994), "The R&D Effectiveness Index: A Metric for Product Development Performance," *Journal of Product Innovation Management*, 11, 213-220.

19. Jaffe, Adam B. (1986), "Technological Opportunity and Spillovers of R&D: Evidence for firms Patents, Profits, and Market Value," *American Economic Review*, (December), 984-1001.

Table 1. Summary of Research Findings
(Categorized using the Tier Metaphor)

Tier 1 Basic Research Explorations	1. Research tourism encourages research spillovers which enhance long-term profitability. 2. Metrics based on <u>all</u> ideas, no matter their source, match R,D&E's incentives with those of the firm. 3. Metrics which reward people for internal ideas lead to (a) too few ideas, (b) excessive research empires, and (c) "not-invented-here" actions and decisions.
Tier 2 Programs to Match or Create Core Technological Competence	1. Metrics must recognize that program decisions differ from decisions on applied projects. 2. Metrics must recognize that the <u>choice</u> of research program is critical and that it is made before most of the scientific, engineering, and process effort is undertaken. 3. Sole emphasis on market-based outcome metrics is counter-productive when <u>choosing</u> research programs. Market-based outcome metrics should be used but given a small relative weight. 4. However, <u>after</u> the program is chosen, R,D&E must encourage the right amount of scientific, engineering, and process effort. This requires <u>effort</u> metrics to balance <u>cost</u> metrics. 5. Traditional metrics, such as publications, citations, patents, citations to patents, and peer review, can serve the role of effort metrics. The best metric systems use a combination of effort metrics and market-outcome metrics.
Tier 3 Applied Projects with or for Business Unit "Customers"	1. Business units have an important say in the choice of applied projects, however, if they have the only say, then they will choose projects that are shorter-term, less risky, and more focused than is best for the firm. 2. Subsidies can be used to adjust for short-termism, risk aversion, and narrow scope. 3. However, to be efficient, the level of subsidy should vary by firm and by project according to the formula, $S = \gamma R \alpha$. 4. Options thinking should be used to measure the value of flexibility in decisions to continue projects. (This will lead the firm to accept more uncertainty.)

Table 2. R&D Metrics Reported by Interviewees (from 6)

	Category	Metric	Most Relevant
Qualitative Judgment	Strategic Goals	Match to organization's strategic objectives	Tier 2
		Scope of the technology	Tier 2
		Effectiveness of a new system	Tier 2
	Quality/Value	Quality of the research	Tiers 1, 2, 3
		Peer review of research	Tiers 2, 3
		Benchmarking comparable research activities	Tiers 2, 3
		Value of top 5 deliverables	Tier 3
	People	Quality of the people	Tier 1
		Managerial involvement	Tiers 2, 3
	Process	Productivity	Tier 3
		Timely response	Tier 3
	Customer	Relevance	Tier 3
Quantitative Measures	Strategic Goals	Counts of innovations	Tier 2
		Patents	Tier 2
		Refereed papers	Tiers 1, 2
		Competitive response	Tier 3
	Quality/Value	Gate success of concepts	Tier 3
		Percent of goal fulfillment	Tiers 1, 2
		Yield = [(quality*opportunity*relevance*leverage)/overhead]*consistency of focus	Tiers 2, 3
	Process	Internal process measures	Tiers 1, 2
		Deliverables delivered	Tier 3
		Fulfillment of technical specifications	Tier 3
		Time for completion	Tier 3
		Speed of getting technology into new products	Tier 3
		Time to market	Tier 3
		Time of response to customer problems	Tier 3
	Customer	Customer satisfaction	Tier 3
		Service quality (customer measure)	Tier 3
		Number of customers who found faults	Tier 3
	Revenues/Costs	Revenue of new product in 3 years/R&D cost	Tier 3
		Percent revenues derived from 3-5 year old	Tier 3
		products	Tier 3
		Gross margin on new products	Tier 3
		Economic value added	Tier 3
		Break even after release	Tiers 2, 3
		Cost of committing further	Tiers 1, 2, 3
		Overhead cost of research	