R&D PERFORMANCE MEASURES THAT ARE LINKED TO STRATEGY

by

C. H. LOCH*
and
U. A. S. TAPPER**

2000/14/TM

* Associate Professor of Operations Management at INSEAD, Boulevard de Constance, 77305, Fontainebleau Cedex, France.

** University of Witwatersrand, Johannesburg, South Africa.

R&D Performance Measures That are Linked to Strategy

Christoph H. Loch,
INSEAD
Fontainebleau, France

U.A. Staffan Tapper
University of Witwatersrand
Johannesburg, South Africa

Abstract

Although most companies recognize the importance of R&D for future success, they often struggle to assess R&D performance. No widely accepted performance measurement system for R&D exists. Performance measurement is particularly difficult for the R&D function because the success of a new product or process can only be assessed with certainty after a long delay. Furthermore, success criteria are not always known because they are strategy dependent and because success drivers are subject to factors outside the control of the R&D function. It is, therefore, difficult to formulate measures that are timely, connected to business goals, and fair. We show how the Research Group of the diamond producer GemStone developed a research performance measurement system that is appropriate for the risky projects typical in R&D, and which supports business strategy.

The process derives operative measures for R&D from the company strategy. It starts top-down, but then facilitates bottom-up formulation of initiatives. In addition, a mixture of output-oriented and process-control measures ensures timeliness and fairness of the system.

In the summer of 1999, the research group of GemStone, a medium sized diamond-mining company, held a workshop on their campus in Winnipeg, Canada. One researcher, an eminent geological physicist, burst out: “I am working on three diamond extraction projects, but you know what? I am making progress on none of them, because I get 35 phone calls per day not only from Technical Services, but also from the mines directly. Because of some breakthrough I did ten years ago, they know my name as the recovery expert, and they all call me when they have trouble. What do you want me to do? Turn them off or continue to become obsolete in my field?” Larry, a geologist, added: “We don’t get any respect in the mines because they think we just tinker with crazy stuff. And honestly, I cannot explain to them how what I do relates to their business, except maybe five years from now. I don’t have any guidance what to work on. Yes, I’m the one who is technically most knowledgeable and I need to make the last call in what I do. But at least the company needs to tell me into which direction they want me to run!”

Research management was aware that they had no good way to communicate the value of research to the organization. The annual research budget was usually last year’s plus inflation (or less, if cost cutting set in), but they could not really justify it, nor could they argue why the company might want to increase its research expense. The rest of the organization did not understand what the research group did, and was, therefore, not in a position to appraise the value adding resulting from its R&D investments. This left a nagging tension with the operating units, an unspoken doubt whether research was free-riding on them.

The Difficulty of R&D Performance Measurement

The GemStone research group’s situation is not so uncommon. As product life cycles shorten and competition heats up, the importance of Research and Development (R&D) is increasing in many industries. Yet, a majority of companies struggles in determining how well the R&D function of their company is *really* doing. If one asks the head of R&D, setbacks are due to higher forces. If one asks marketing, the setbacks are due to R&D incompetence and the successes due to good preparation by marketing and to a lucky easy ride. Wouldn’t it be nice if we could evaluate R&D with financial criteria? Evidence shows that this backfires. Financial measures of

R&D performance are often used poorly and inappropriately. A recent survey shows that companies using financial measures in their R&D funding decisions perform *worse* than companies that don't (Cooper *et al.* 1999).

To see the danger of financial measures, consider the example of the VP of R&D in a white goods manufacturer. His company had introduced the Economic Value Added (EVA) framework, cascaded down all the way to individual R&D projects. What does EVA at the project level boil down to? It measures returns in excess of capital cost, which corresponds to nothing else than a net present value (NPV) discounted at the aggregate cost of capital! Thus, every department head had to fulfill certain NPV goals. The question became: do I choose the innovative project that could pay its investment a hundred times over, but it's likely to fail? Or do I put my budget into the process improvement that will almost certainly save $250K for an investment of $200K, yielding a good return of 25% in one year? As a result, a significant shift occurred to safe incremental projects of the second type, because only with those could managers be confident, project by project, to make their numbers. The organization lost much of its innovativeness.

An additional problem is that managers typically hope for general measures that they can benchmark against other companies. For example, measures that are widely cited from "role model companies" are time-to-market ("Toyota can develop a car in 18 months"), the new product sales ratio ("3M derives 35% of sales from products less than 3 years old"), or R&D intensity ("Pfizer has the highest R&D ratio with 19% of sales"). Alas, transferring such ratios from other companies to your own is dangerous; another recent study has shown that they predict success only for dominant companies in slowly growing markets with long product life cycles. In other industries, they are meaningless (Terwiesch, Loch & Niederkofler 1998).

These shortcomings prompted Kaplan and Norton to develop their balanced scorecard, which explicitly recognizes that additional operational measures (customer-related, process-related and learning-related) should be derived from the company's strategy (see Figure 1).

However, the balanced score card does not solve our problem of how to evaluate the performance of the R&D function. It is designed for measurement at the business unit level, too aggregated a level to speak directly to R&D. In theory, R&D is just one of the internal business processes in the scorecard (on the right in Figure 1), and Kaplan and Norton explain how to derive process measures using the examples of manufacturing and sales. The problem is that R&D measures cannot be formulated according to the same logic because R&D exhibits two critical complications: *uncertainty resulting in causal ambiguity*, and *long time lags*.

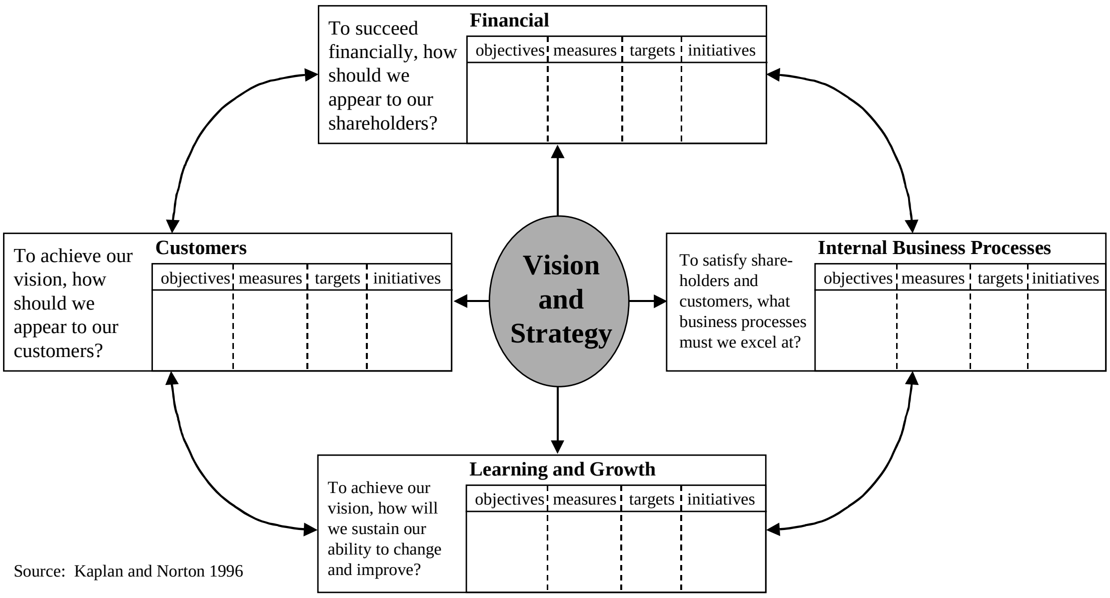


Figure 1: Kaplan and Norton's Balanced Score Card

In discussions we have had with R&D managers, *over a third of them believe that it is impossible to measure R&D adequately*. Bill Hewlett, co-founder of HP, once remarked: "What you cannot measure, you cannot manage. What gets measured gets done." Does this mean that R&D cannot be managed? We do not think so. The challenge is to develop detailed and tailored measures for R&D, measures appropriate under ambiguity and time lags between action and result, and which can serve R&D employees as helpful guides in fulfilling their contribution to the company's business. In this article, we show that this is possible by carefully breaking down the business mission into operational activities.

Example: Taking a Trip in a Car

Imagine you want to measure your performance during the last car trip that you took. Figure 2 illustrates that there are actually three types of measures that you need to look at. First, *operational control* relates to real-time information allowing immediate reactions to complete the trip successfully and efficiently: direction (staying on the road and avoiding oncoming traffic), speed (making the next curve, and avoiding traffic tickets), engine revolutions (to shift gears), remaining fuel (not running out), and obstacles in the way (avoiding crashes). Second, *input measures* indicate the amount of resources spent on the trip (such as time, fuel or money, or scarce personnel).

- cost
- personnel
- time
- management attention

- Follow winding road (stay on course)
- Keep safe distance from the car ahead of you
- Shift when engine revolutions go too high/too low
- Stop at red lights/ crossovers
- Refuel when tank is low

- Evaluate driver
  - distance, time, avg. speed (distance/time)
  - fuel consumption, fuel efficiency (miles/gallon)
  - safety (accidents, “near misses”)
- Learn for Future
  - where are traffic jams?
  - which roads are good?
  - is the car reliable?

Figure 2: Performance Measures for Taking a Car Trip

Third, *output measures* serve a double purpose. *Evaluation* determines whether the driver has done a quality job. This refers to the trip’s results (was it safe, was the goal reached) or to efficiency, and often the ratio of outputs to inputs (such as speed and miles/gallon achieved). It is important to note that these evaluation measures should reflect not only one goal but also the relevant *trade-offs among goals*. “What you measure is what you get.” If only speed is measured, the driver will go fast, but burn excessive fuel and run the risk of traffic tickets or accidents. If only cost is measured, the driver may go so slow that the target arrival time is missed. A combination of cost and speed measures can push the driver’s behavior into the required compromises for achieving the desired output.

The second purpose of tracking output measures is *learning*, for example, an analysis after the trip which route was congested, and where the roads were bad. Thus, the driver can take a better route next time, improving performance.

So, now we understand how to measure a car trip: what measures support execution, and what measures allow to evaluate the driver and to learn. Or do we? Imagine *three different* car trips – a truck journey to deliver some commodity to a factory, a formula-one race, and a Sunday drive across country with the family. Does the above set of measures fairly evaluate these three trips? The reader will quickly realize that the answer must be no – their purposes are too different, and they have different sets of stakeholders. In other words, measures are meaningless without considering strategy. Figure 3 summarizes this.

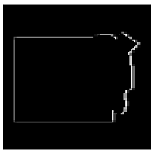
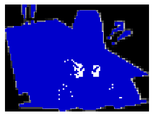
	 **Commodity delivery**	 **Race**	 **Sunday Drive**
Key stakeholder	Truck owner	Race team and driver	Driver and family
Trip mission	Deliver reliably, on time and at low cost	Complete distance at minimum time	Fill time, have fun
Operational control measures (in addition to steering and breaking)	Distance made vs. plan Respect speed limits Obstacles (traffic, roads blocked for trucks)	Speed Engine revolutions (rpm) Position in race Tire condition, weather	Distance from home Obstacles (traffic jams) Proximity of restaurant Weather
Evaluation measures	Cost On-time performance Safety (no accidents, tickets)	Lap speed Ability to overtake Safety (no accidents) Strategy (e.g., refueling)	Attractiveness of road and places visited Flexibility to stop Safety (accidents)
Learning measures	Route conditions Traffic conditions Natural stopping points Routes allowing return freight	Mechanical performance Mechanical failure reasons Strategy	Traffic conditions Attractiveness of route Conversation during trip

Figure 3: Evaluating the Performance of Different Car Trips

The key stakeholder of the truck trip is the owner, who wants (a) no accidents and (b) reliable on-time delivery. In the race, both the driver (pursuing his career) and the race team (pursuing continuing support from its sponsors) are stakeholders, and the fastest time is the thing that matters by far the most. Safety is a concern, but not as much as for the truck: winning requires taking risks. For the Sunday family ride, time is no issue at all (except for being home for dinner). What matters is enjoying the process of the ride, and the family (who may want to go slow to see the scenery) has as much to say as the driver (who may want to go fast because that's fun). Driven by the different goals for the three trips, the operational, evaluation and learning measures must differ accordingly. If the driver behaves according to the performance measures for a race, but the family thinks they are on a Sunday drive, the result will be a family crisis!

This illustrates the first important principle: performance measures must help you to implement and monitor your specific strategy. Generic measures and generic benchmarking are meaningless. For example, BMW takes longer to develop a car than Toyota, but that's consistent with their strategy of pushing the envelope each time. Or, the R&D intensity (% of sales) is of course lower for a generic drug-marketing firm than for the top research firms, as they compete differently! This is how far Kaplan and Norton's Balanced Score Card can take us, as it offers a measurement framework at the strategic level of the business unit. However, R&D has several additional requirements for successful performance measures, which are unique and not addressed in the Balanced Scorecard.

Meeting the R&D Performance Measurement Challenge

Let us go back to the required characteristics of useful measures. Operational measures must be directly connected to the operational sub-goals of the trip. One operational sub-goal is "stay on the road, without accident." The distance to the car ahead must be estimated accurately in order to prevent accidents; measuring the distance to the trees off the road would not be connected to the trip and thus useless. In addition, their feedback must be immediate; for example, the driver needs to react instantly to the car ahead breaking. If these two requirements are not fulfilled, the trip will end prematurely, and badly.

The same requirements hold for output measures in their relation to the trip's mission. They must be *timely*, as neither driver evaluation nor learning are possible if the feedback delay is too long: finding at the end of the life of the car that the engine could have lasted 10,000 more miles is too late to evaluate the driver's performance of this trip (he will have forgotten the specific trip, or may no longer be there). Nor does it allow learning, as the environment may have changed, making the causes of the original trip obsolete (e.g., different roads may now be prone to traffic jams)[1].

In addition, a *causal mechanism* between actions and the measured construct must be known, and this causal connection should not be obscured by other factors ("statistical noise"). Suppose that in the car race, an engine with a newly developed technology is used, which is not yet reliable. The driver may be the best in the world in still come in last (if the engine does not perform) or drop out with engine failure. This weakens the connection between the driver's behavior and the race's success. Looking at success alone will not allow telling how good the driver is. Similarly, if the Sunday trip occurs in an area that the driver knows, the family can hold him responsible if the trip was not fun, but if they do the trip in an unknown country, he cannot promise anything. The family can evaluate him only on trying.

Timeliness and causal connection to the mission are the two areas where R&D poses unique challenges to performance measurement:

- *Long delays.* It has long been known that it takes a long time, often up to half of the life of a product after it has been introduced in the market, until one knows with acceptable precision how profitable that product really is (Beardsley and Mansfield 1978).
- *Causal ambiguity and Uncertainty.* Many factors may play a role in the success of a new product or technology, and most of them interact. Thus, it is often not clear which ones are connected to the overall success. Moreover, significant factors are often unknown and cannot be predicted by the product development team: a competitor move, a demographic change in the market, a political event in the country of the target market. The team has done everything right, but the product failed. In this sense, novel (and promising) projects are *risky*, because not every promising effort produces a business-relevant output. A good researcher is not necessarily the one who always produces some business result (the results may be very conservative and mediocre), but the one who recognizes the potential of

[1] This is well described in Meyer 1993.

an effort, while being willing to abandon it if he learns that the potential does not materialize.

The delays and uncertainty pose further requirements for R&D performance measures: *Cascade* long-term strategic success into short-term proxy measures that are close to action, and *Measure According to Risk:* at the project level, measure output for plannable efforts, but process for risky projects. At the aggregate level of the R&D function as a whole, measure output.

Cascading Measures from Strategy

The essential questions of linking technology performance and strategy are shown in Figure 4: what is the business strategy, what does it need from R&D (top-down), and what can R&D propose to enhance the business strategy (bottom-up)? Within R&D, the same top-down and bottom-up questions can be asked for new product or process development and the underlying technology capabilities.

As Figure 4 summarizes, business strategy can succinctly be characterized by the five questions at the top of Figure 4: *what* do we sell, to *whom*, *how* (with what core competences), *why* (what is the competitive advantage, or value proposition to the customers), and what are major *threats* in the environment (Markides 1999). The strategy must also identify tradeoffs and priorities among conflicting goals (e.g., emphasis across market segments).

A technology strategy (bottom of Figure 4) must specify what technologies to master, how this portfolio of efforts relates to the business: what products and segments it addresses, and what it contributes to competitive advantage and to hedging against environmental threats. The technology strategy also has a *how*: with what timing, and what risk profile are the products/processes delivered, and how does this program fit into the available set of resources (Porter 1985, Roussel *et al.* 1991, Cooper *et al*. 1997). Performance measures can then specifically examine whether the thus formulated strategic mission is met (and if not, why).

The technology strategy needs to be *cascaded down to increasingly operational measures*, at the department level, to which the R&D employees can relate. As

Figure 4 shows, the R&D measures should be a "mirror image" of the strategy evaluation: they should evaluate the technology strategy along the same questions used by top management to evaluate the business strategy, broken down to assess the *pros and cons* (that is, trade-offs) of R&D activities for the business. At the same time, this framework of measures allows researchers to express ideas in a way that relates to the overall strategy, making explicit when a new idea is not aligned with strategy, giving the researcher a way to justify why he/she wants to deviate. The set of measures can become a vehicle of structured strategy modification.

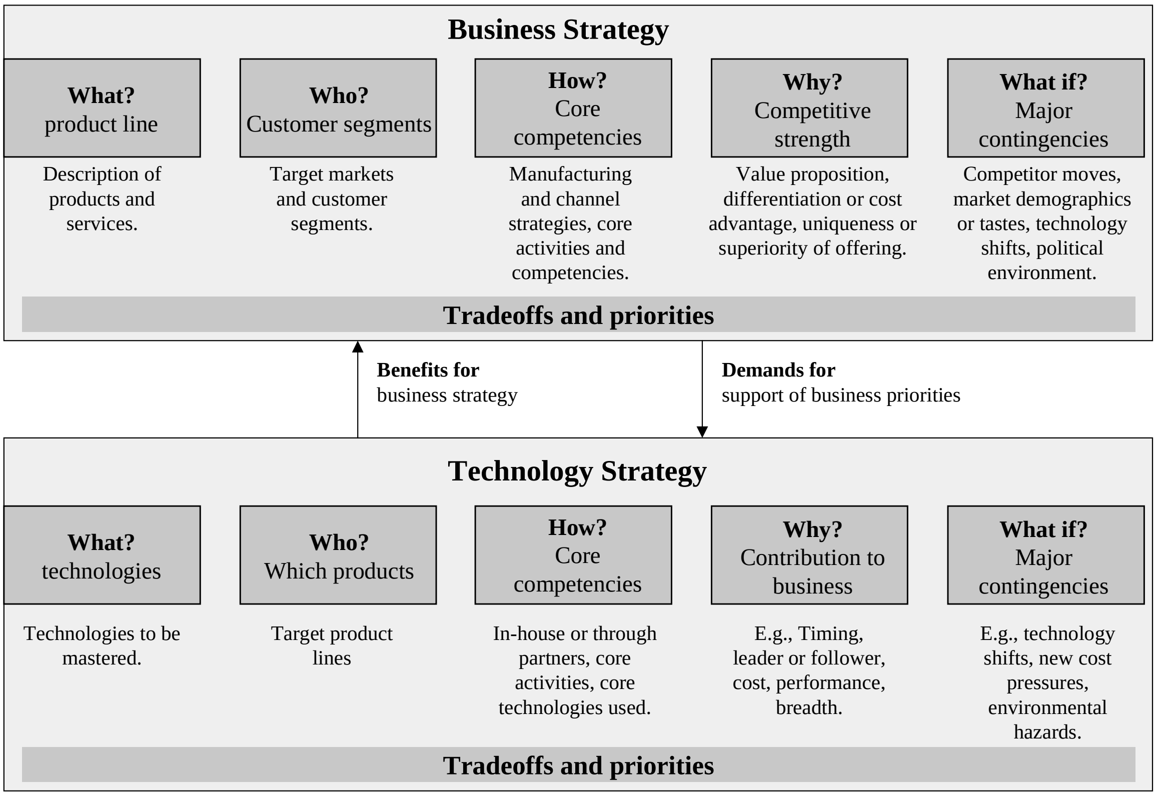


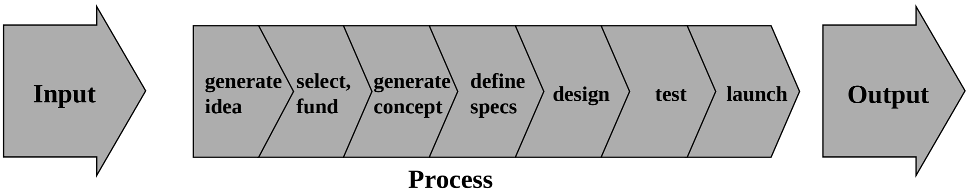
Figure 4: Linking Business and Technology Strategy

The Balanced Scorecard measures of a technology strategy (financial success, growth, and customer satisfaction) fit under the "why: contribution to business" in the technology strategy at the bottom of Figure 4. Our framework gives an operational guideline, missing in the Balanced Scorecard, to the manager as to how the cascading down to the R&D department can be performed.

Measure According to Risk

At the project level, measure output or process depending on the project uncertainty. For routine projects, the project team can influence outputs (deliverables), and can thus be held responsible for them. In such a case, one may even be able to calculate even efficiency figures (output per input), such as “revenue per R&D employee”, which some companies apply. Figure 5 lists some widely used project input/output and process measures.

In a project with high uncertainty, however (technical or market uncertainty), output measures are not under the control of the researcher and thus not appropriate as evaluative or operational measures. Therefore, the individual researcher (also) should be measured based on the quality of the research *process* he/she follows.



- **cost**
- **personnel**
- **time**
- **management attention**

- **Link to strategy, vision**
- **top management support**
- **cross-functional teams**
- **project management methodology**
- **intermediate results**
- **strong project manager**
- **champion**
- **customer involvement**
- **intensive communication**
- **quality of personnel**
- **(peer review)**
- **(…)**

- **product quality: design, performance, conformance**
- **manufacturing, delivery or life cycle cost**
- **time-to-market**
- **R&D cost revenues**
- **profits, ROI**
- **market share**
- **technical goals**
- **capability building**

Figure 5: A “Laundry List” of Typical Project Output and Process Measures

However, even if process measures are preferable in some projects, output must be measured and tracked *somewhere* in the organization. To return to the car racing example, the driver may not be responsible for the result of a single race (because of unreliable technology and the inherent riskiness of racing), but the performance over a whole season does reflect on his abilities. And the team manager is definitely held responsible for the result of the season, as he has several cars running and makes decisions about the technology used (thus determining the level of risk taken).

Analogously, risks of individual projects can be “averaged out” at the aggregate level of the research department, and output is controllable (by improving processes and by allocating resources to a good portfolio). While business payoffs can only be tracked historically and in hindsight (because of the long feedback times), there exist proxy measures of good output, which can be estimated immediately at the end of all projects, and then aggregated.

Measures that are widely used are, for example, the new sales ratio (sales of "new" products as a% of total sales), the cost savings ratio, profits from new products (as a % of total products), the R&D yield (profits from new products over total R&D costs), the number of products launched per year, the % of projects launched that are successful, the % of products that are 1st to market, or the average age of products currently in the product line relative to the industry life cycle. More elaborate concepts have been proposed, for example, the "R&D effectiveness", an extended version of the R&D yield.

In summary, we have identified the required characteristics of R&D performance measures; they are summarized in Figure 6. We now show how we implemented a system with these characteristics at the research group of GemStone.

Summary: Characteristics of Effective R&D Performance Measures

- Link to Strategy: cascading from business strategy through technology strategy down to department strategy, and feed insights back up. At department and project level, measures are proxies for the higher-level strategic goals.
- The R&D measures and the strategic portfolio of R&D programs are **mirror images** of one another: the programs execute strategy, and the measures monitor status and progress.
- The R&D measures capture key business trade-offs inherent in making R&D decisions in the company
- The proxies should be:
  - timely,
  - influenced by behavior ,
  - for low-risk projects: measure the *output* (result),
  - for high-risk projects: measure the quality of the *process* used,
  - at the level of the research group: measure group output versus the strategic goals and their shorter-term proxies.

Figure 6: Summary of R&D Performance Measure Characteristics

Deriving Performance Measures at GemStone's Research

We demonstrate the process of linking R&D performance measures to the strategic mission on the example of GemStone, a small diamond producer[2]. Digging diamonds out of the ground may seem simple, but it turns out that it requires sophisticated

[2] See the INSEAD case "GemStone". The company name is disguised for confidentiality reasons.

technology to find the existing diamonds at acceptable cost. We concentrate on research, rather than development, because research exhibits the problems of long time lags and uncertainty particularly strongly, and thus represents a good test application of the above-developed principles.

The Technology of Diamond Production

Diamonds are formed out of carbon that has been subjected to extremely high pressure and heat. In the early days of diamond mining, until the 1860s, gems had been washed out from surface deposits. But these easily exploited sources were soon exhausted. Today, diamonds are mined from eroded *kimberlite pipes*, formations of crushed rocks shaped like ice-cream cones that volcanic activity has thrust to the earth's surface from depths that can exceed 150 km. Mining is very capital-intensive – typically, over 1,000 tons of rock must be excavated and processed in order to extract 100 grams of diamonds of sufficient size and quality.

The steps of the diamond production are: exploration, ore evaluation, mining, and extraction (see Figure 7). Exploration performs aerial photographic surveys and has teams of geologists in the field, who collect rock samples from geologically promising ground formations to identify kimberlite deposits. Sometimes, the kimberlite is deep under ground and cannot be discovered directly, but only via so-called *indicator minerals*, other types of rock that statistically tend to be found close to kimberlite. There is intense competition among diamond producers in buying prospecting rights in promising areas and securing mining licenses whenever a feasible mine has been identified. Maintaining the prospecting right in an area is expensive (the local authorities have caught on to the revenue potential from mines). Very few kimberlite deposits turn out productive -- of the several thousand kimberlite occurrences known in 1990, only about 50 are considered commercially viable. Therefore, fast feedback from explorations is deemed critical.

The ore evaluation group's role is to "understand what is in the ground and couple it to the market". Once a deposit has been identified, this group estimates its richness and the grade of the diamonds it contains, based on geological models they have developed over the last 30 years. They translate this data into a "value estimate" for the potential mine and thus determine the feasibility of exploiting it. If a deposit is judged feasible for exploitation, it is upgraded to a "reserve." During later mining, the ore evaluation group regularly monitors actual production, looking for variances to their predictions, which helps both to identify possible inefficiencies in operations and to continuously upgrade their models.

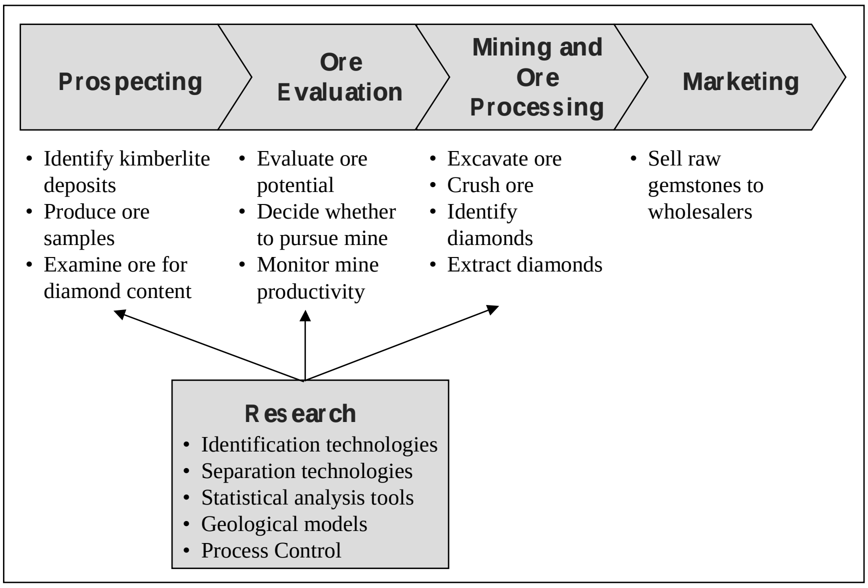


Figure 7: GemStone's Business System

Diamond extraction has traditionally been done by crushing the rocks to small sizes, and, following various concentration processes including density separation, passing the rock fragments on a conveyor past personnel who picked out the diamonds by visual inspection. This process poses several problems. First, a balance has to be struck in the size to which the rocks are crushed. If the rocks remain too large, diamonds remain hidden in them and are lost in the process. If crushing is too fine, large diamonds are damaged or even destroyed. As the value of a gem grows steeply with size, one tends to err toward larger rocks. In addition, large gems represent an almost irresistible temptation for the inspection personnel. Tight security systems have been installed to prevent theft. One manager estimates that currently, 10-20% of value are lost due to overlooking them, damage or theft.

Cascading Strategy to the Research Mission

We can concisely summarize GemStone's business strategy by answering the five questions from Figure 4. Figure 8 refines Figure 4 to show the cascading of GemStone's business, development and research strategies[3].

[3] The strategy is simplified, in order to focus and clarify exposition, and to preserve confidentiality.

	What?	Who?	How?	Why?	What if?
Business Strategy	Raw gemstones	Diamond cutters and wholesalers	Core Processes: • Prospecting • Ore evaluation • Mining and ore processing • Partner producers	• Price • Quality (purity, size, brightness) • Breadth of stone assortment	• Additional producers (e.g., Mongolia, Caucasus) • New deposit areas, e.g., deep sea, Antarctica • Emergence of synthetic diamonds as gemstones
Development Strategy	• No "products" (stones are found) • *Processes:* - equipment - techn. services (e.g., control & improvement) - analysis & knowledge	Core processes • Prospecting • Ore evaluation • Ore processing • (mining equipment bought externally)	*Organization:* • Equipment development & manufctg. • Technical services • Research *Process:* • stage gate process	• Kimberlite detection rate • Process cost • Process yield (% gems discovered, % gems destroyed) • Technology frontier awareness • Ability to mine in new areas • Technology reputation to attract partners	• Emergence of synthetic diamonds as gemstones (technology shift)
Research Strategy	• Technology demonstrations (reports, recipes, prototypes) • Breakthrough concepts • Customer support • Technology knowledge • Technology reputation	Outside R&D: • Prospecting • Ore evaluation • Ore processing Within R&D: • Development • Techn. services	• Project management • Focus on core areas • External partners (labs, universities) • Publication and peer review • "Customer" training	• Creation of value opportunities • Enhancing "customer" ability to perform their mission more effectively	• Technology shift

Figure 8: GemStone's Business, Development and Research Strategy

GemStone produces essentially a commodity product, using the core processes discussed above. Diamonds are sold to cutters based on quality, price, and the completeness of diamond varieties one has to offer[4]. There are several potential threats to the industry structure (artificial diamonds, the emergence of new producers, and new deposit areas contemplated if not yet utilized). GemStone is attempting to build an alliance of producers to strengthen market position.

As GemStone does not *produce* diamonds, but *finds* them, innovation concerns only new processes, not new products. The R&D organization comprises (process) *research* and *development.*

In this environment where technology has no influence on the quality of the product, key contributions from new process development concern business process *costs* and *yields*, in the core processes of prospecting, ore evaluation, and ore processing. R&D

[4] The fact that diamond producers attempt to "de-commoditize" their product with marketing efforts aimed directly at the end customer is beyond the scope of this article.

is also responsible to serve as a knowledge repository of new technologies and how they may affect the business, specifically to prepare the operating units for new areas of mining expected to emerge soon: under the sea and arctic. In addition, top management has formulated the mission to develop a technology leader reputation, in order to be more attractive as a potential partner to other producers. Corresponding to this structure of business needs, the development organization has two arms: an equipment development and manufacturing unit, and a technical services unit (auditing and improving the existing processes in the operative units).

Research has an annual budget corresponding to 100 person years (PYs). The 50 most highly qualified people have Master's and Ph.D. level degrees and include physicists, electronics engineers, software engineers, mechanical engineers, metallurgists, control and instrumentation experts, and mining engineers. In addition, research maintains an extensive network with external R&D contractors and universities. The research organization has the responsibility to create new business opportunities for their "customers," and to provide them with technologies helping them to perform their respective missions more effectively. Research "customers" are the operating units (prospecting, ore evaluation, and ore processing) as well as to their direct partners within R&D, namely equipment development and technical services.

Research proposes to create this value with four types of outputs:

- Technology demonstrations, that is, proof of feasibility and potential. This can be done with working (hardware) prototypes, process recipes, or with technical reports proving a principle.
- Breakthrough new concepts (e.g., identifying large diamonds in the rock with hard X-ray radiation)
- Being a "knowledge repository" for the whole company about all technical aspects of diamond production, advising on high level decisions impacted by technology and on current trends. This includes training of personnel (e.g., for technical services).
- Providing an external technology reputation, for example, through conferences or publications.

It is a key feature of this effort that it was *the researchers themselves* (research program managers) who formulated their strategy, facilitated by the manager of research and the authors. Treating them as the professional experts who can judge their work best is an important part of getting buy-in and encouraging creativity and initiative. They decided to focus on core areas of expertise[5] and to emphasize collaboration with external partners (universities and private labs), also strengthening project management and start evaluating external publications. Finally, they proposed to improve the cooperation with their "customers" by offering technical training, particularly for technical services.

The research group had not yet formalized a strategic program portfolio. Projects had been started based on ad-hoc initiatives, e.g. upon the initiative of powerful or outspoken operations managers (in prospecting, ore evaluation and ore processing). As a result, the researchers felt that they were working piecemeal, driven externally rather than based on their own judgments about technology trends, and were spread over too many activities to be productive. In addition, they felt that they could not explain to the rest of the organization in terms understandable to them what they did. The above analysis allowed the research group to

(a.) *formulate a collection of programs* (a portfolio) that had a clear rationale versus the strategic goals of the company (top-down), but at the same time allowed the researchers to formulate their own ideas of how they might be able to contribute to the high level goals (bottom-up).

(b.) *formulate a set of performance measures* that reflected the specific characteristics both of the strategic mission and of the work to be performed, and was adjusted to the risks so they could be seen by the employees as constructive and fair rather than as an instrument of control.

The research group came to understand that performance measures are the *mirror image* of the activity portfolio (see Figure 8). A strategic portfolio allows formulating a set of activities that embodies business strategy, while performance measures allow monitoring the progress and the value produced. In this article, we concentrate on the measuring side; for excellent descriptions of the principles of portfolio management

[5] Strategy and areas of expertise have been simplified and shortened to protect sensitive information.

see Roussel *et al.* 1991, or Cooper *et al.* 1997. GemStone's research group began by formulating a set of measures first; the formulation of an activity portfolio has begun in parallel and is in progress.

Developing Research Performance Measures

As an output of this effort, the group developed a set of research performance measures, which is summarized in Figure 9. The research manager is responsible for an overall adequate "output." This output is measured only partially in financial terms because financial measures are too uncertain and far off in time. Rather, the measures are *operational and can be influenced* by the researchers' activities. The measures also *capture the relevant tradeoffs*. For example, the number of equipment technologies is counted, but also their innovativeness and the maturity of their transfer (the latter at the project level, which should be aggregated up and be applied at the group level as well). Finally, the measures *capture the strategic needs* from Figure 8, addressing technology innovations, breakthroughs, support, knowledge creation, and external reputation.

	New technologies	**Customer support**	**Knowledge repository**
Output measures (group level) *(by customer: exploration, mining, ...)*	• # of significant innovations delivered • impact of the technologies delivered (qualitative estimation by customer, follow-up to learn) • Market potential of innovations in $ • # of presentation to *external* customers to which research contributes	• customer satisfaction index • Response time to queries • % of support requests fulfilled • # of training sessions signed off by customer and delivered • # of problem analysis reports requested and delivered	• number of requested handbooks" published and delivered • quality of research program homepage in intranet: # of hits • external reputation: # of external publications, patents, and their impact
Output measures (project level)	• level of prototype maturity (e.g., no. of major technology revisions after hand-over to development) • quality of documentation to development	• response time • request fulfillment	• successful study completion • Clear go/ no go decision
Process measures (project level)	• professional schedule and budget planning and control • conscientious management of uncertainty (e.g., risk reduction assessment) • professional documentation, peer review • communication within research and with customers • use of external knowledge, cooperation with partners	• quality of interaction with the technical services requestor	• completeness of literature surveys • documentation • clarity and quality of conclusions in technology assessments (e.g., by peer review) • "project management" of writing handbooks and assessments

Figure 9: Performance Measures for GemStone's Research Group

New technologies are, by definition, more uncertain and risky than support activities for internal customers. Knowledge repository activities (right column in Figure 9) are not risky in the same sense as the attempt to create new technologies, but they are intangible, it is hard to measure the direct usefulness of a piece of knowledge stored and made available. Therefore, *process measures* are more important for new technology and knowledge repository projects than for support.

The measurement framework includes both output and process measures at the project level. The research manager and the researcher can set the appropriate balance on a *case-by-case* basis (by choosing appropriate measures from the framework). In very risky projects, output should be used as an *upside only*, that is, a reward if the researcher manages a success, but no punishment when the project fails (at the same time, the researcher is responsible for professionally conducting the project – sloppy execution is not to be encouraged). Thus, it can be ensured that the measures *motivate* rather than dulling or causing a temptation to cheat.

Benefits of the Performance Measures for the Research Group

Empowerment and Creativity

A key aspect of this initiative has been that R&D employees give *themselves* a process for diagnosis and improvement rather than having a system imposed. This is similar to what Adler *et al.* (1999) call an "enabling bureaucracy," where a structure (here: for measurement of contribution) is introduced, but *in an atmosphere of trust, and with substantial input from the employees*, who get the chance to set "best practice" targets. They are treated, and thus are encouraged to behave, as professionals rather than order-takers.

Remember the race driver from our introductory example. He will scoff at any attempt of an outsider to tell him how to drive (he is, after all, one of the best!). However, he will be interested in suggestions about "standard operating procedures" (e.g., shifting, staying in the wind shadow, overtaking, refueling, ...) that may help him to become more competitive. HE will decide which to adopt, but he will be grateful for someone suggesting success statistics ("with an *x*-second acceleration advantage you need *y* hundred yards of straight track to successfully overtake"). The situation is comparable in the sense that researchers can never be fully monitored, as their work is too idiosyncratic and complex. Of course, research output is also harder to measure than the race driver's success. But just like a race driver, researchers can (and do at GemStone) give themselves "best practice" procedures that increase their chance of success.

It is absolutely essential that management credibly maintains the use of the measurement framework for transparency and fairness; once the trust is lost, the employees will effectively prevent management from looking over their shoulder, and the effort will collapse.

Bottom-Up and Top-Down Integration

The framework of measures allows *bottom-up proposal of ideas* within a structure. The framework *educates researchers about the business needs* of the organization and helps them to think strategically. Although the framework initially must be derived top-down (as we illustrate above), it then gives researchers a way to think about and categorize new ideas: they can recognize and justify a new idea as fitting a certain set of activities and measures, or express in what sense a new idea does not fit the current strategy, but has a potential that could complement or modify the strategy. Thus, the quality of communication between top management and researchers can be enhanced.

Thus, the framework of measures becomes a *communication device*. It allows researchers to more easily communicate the contributions of research to top management as well as the rest of the organization. This is of value for the motivation of research employees. In addition, the transparency resulting from a research performance measurement system may also improve the relationships with management and facilitate proposals for project funding.

Project Prioritization

The framework allows better prioritization of efforts – it is the mirror image of the activity portfolio. Once a framework is in place for estimating the value and contribution to strategy of research projects, a better intuition can be built for the difficult and non-quantifiable process of project evaluation. This facilitates the setting of priorities and the construction of an effective portfolio of R&D efforts. A discussion of the research portfolio construction is beyond the scope of this article; it is ongoing at GemStone.

Benchmarking

The framework tells you *where you can benchmark*. Benchmarking creates insights and learning only on measures that are used for similar reasons in both the benchmarked and the benchmarking organization. For example, a company that pursues an aggressive policy of generating revenues from licenses must have different goals for patent generation than GemStone, who only wants to set a signal of being competent in technology. Benchmarking of patents between these two companies would be confusing rather than helpful. What may be comparable are general “affordability” estimates in commodity industries: “if you spend more on R&D than others in your industry (e.g., as a % of sales), do you have an argument why you can derive more benefits from this extra investment than the others?” Beyond such general issues, benchmarks should only be used when the strategic context and goal of the compared measure is understood.

Summary

The right R&D performance measures cannot be scientifically derived. However, they can be designed professionally, with an explicit understanding of how they support the strategy, and considering employees’ normal risk aversion with respect to measures. This paper offers an operational guideline to produce a measurement system that is fair (and thus has a chance of being accepted) and motivates researchers to pursue company strategy.

At GemStone, researchers had wanted to move beyond research and develop and sell the machines they had invented, because this was how one received recognition (i.e., that’s what was measured). The new system allows researcher to focus on their unique role within the company. This has resulted in a much-improved understanding of the role of R&D within the corporation. The system guides the research efforts to be aligned with strategy, and because senior executives were part of the process of developing the measures there is an acceptance that the intangible output of R&D

does indeed further company objectives. Last, but not least, the focus on the new measures has resulted in drastic productivity gains in a short period of time.

References

Adler, P. S., P. Goldoftas, and D. L. Levine. 1999. "Flexibility Versus Efficiency? A Case Study of Model Changeovers in the Toyota Production System." *Organizational Science* 10 (1), 43 – 68.

Beardsley, G. and E. Mansfield. 1978. "A Note on the Accuracy of Industrial Forecasts of the Profitability of New Products and Processes." *Journal of Business* 51 (1), 127 – 135.

Cooper, R. G., S. J. Edgett, and E. J. Kleinschmidt. 1997. *Portfolio Management for New Products.* McMasters University, Ontario, Canada.

Cooper, R. G., S. J. Edgett, and E. J. Kleinschmidt. 1999. "New Product Portfolio Management: Practices and Performance." *Journal of Product Innovation Management* 16, 333 - 351.

Kaplan, R. S., and D. P. Norton. 1996. *The Balanced Scorecard.* Boston, Mass.: Harvard Business School Press.

Markides, C. C. 1999. "A Dynamic View of Strategy." *Sloan Management Review*, Spring, 55 – 63.

Meyer, C. 1993. "Implementation Dynamics and Measures," Chapter 9 in: *Fast Cycle Time.* New York: The Free Press.

Porter, M. E. 1985. "Technology and Competitive Advantage." Chapter 5 in: *Competitive Advantage.* New York: The Free Press.

Roussel, P. A., K. M. Saad, and T. J. Erickson. 1991. *3rd Generation R&D.* Boston, Mass.: Harvard Business School Press.

Terwiesch, C., C. H. Loch, and M. Niederkofler. 1998. "When Product Development Performance Makes a Difference: A Statistical Analysis in the Electronics Industry." *Journal of Product Innovation Management* 15 (1), 3 - 15.