R&D performance measurement: more than choosing a set of metrics

Inge C. Kerssens-van Drongelen and Jan Bilderbeek

University of Twente, The Netherlands

In this article the results are presented of an empirical study focusing on the effectiveness of R&D performance measurement practices in the Netherlands. First, a theoretical examination of the subject 'R&D performance measurement' is given within the context of performance control. A distinction is made between feedback and feed forward control and between the R&D function and the R&D organisation. Subsequently, a description is given of the current practices of R&D performance measurement in terms of measurement purposes, metrics, measurement techniques, norms setting, etc. Furthermore, the influence of contingencies on measurement system design is explored. The data for this research were gathered by means of a survey and nine in-depth interviews. Generally, managers evaluate their measurement systems as being quite valuable, having a positive impact on performance. The findings described in this paper can be used as references for managers to benchmark their R&D measurement procedures: in this respect we distinguish highly and hardly effective measurement procedures. The most important characteristic that seems to distinguish the most effective systems from the less effective ones is customer focus.

Introduction

Once, R&D was considered to be a unique, creative and unstructured process that was difficult, if not impossible, to control. The control techniques used in other business functions were considered inappropriate for the R&D function because control was supposed to harm creativity, and because of the uncertainty of R&D outcome (Roussel *et al.*, 1991). Therefore, control was limited to setting budgets and periodical peer-reviews focusing on technological achievements (Roussel *et al.*, 1991). But, in recent years, changing business environments have challenged companies to improve their R&D processes in terms of effectiveness and efficiency, and senior managers' attention is now focused on R&D's contribution to competitive advantage (Wheelwright and Clark, 1992; de Weerd-Nederhof *et al.*, 1994). Although managers still acknowledge that R&D processes have several characteristics that differentiate them from other processes, they no longer accept that this means that they are unmanageable. As a result, there is a growing acceptance of the need to control R&D processes and, as part of this, to measure R&D performance (Francis, 1992, Schumann *et al.*, 1995).

Cooper and Kleinschmidt (1995) and Griffin (1997) show that the best performing companies already apply explicit performance measurement techniques to R&D. Unfortunately, these authors give little insight as to *how* these companies are measuring, and we did not find many other studies describing current practices in R&D performance measurement. Exceptions include company reports by Patterson (1983), Kuwahara and Takeda (1990), Robb (1991), Francis (1992), Foster (1996), and a few studies reporting metrics as used in practice (Moser, 1985; Griffin and Page, 1993). The objectives of measurement, frequency and timing, measurement techniques as used in practice, etc., have received little attention. Furthermore, we perceived a gap between the methods, metrics, etc. proposed in theory and those we observed in use in companies. Thus, we initiated a research project focusing on current R&D measurement practice in large and medium-sized Dutch companies, consisting of two parts: a survey and a series of in-depth interviews with several R&D managers. The aims of this study were threefold. First we wanted to get an overview of current practice. Secondly, we were interested to find out whether, and if so in which aspects, measurement system design is contextual. Finally, we wanted to gain an insight into what differentiates effective from less effective measurement systems.

Before discussing our findings, we will briefly present the theoretical model of R&D performance measurement we used to structure and to analyse the empirical data.

Theoretical framework

The current interest in R&D control, and more specifically in R&D performance measurement, is reflected in articles having titles and abstracts featuring words such as: effectiveness, performance, success, control, monitoring, assessment, measurement, benchmarking, auditing, evaluation (e.g. De Bandt, 1995; Brown and Eisenhardt, 1995; Cooper and Kleinschmidt, 1995; Hultink and Robben, 1995; Schumann *et al*., 1995; Chiesa *et al*., 1996; Loch *et al*., 1996; Sivathanu Pillai and Srinivasa Rao, 1996). These words are often used as synonyms. However, this can be misleading when trying to interpret and compare the proposed concepts, because the *purposes* and the *subjects* of the measurement efforts, and the *context* for which these concepts are suited, can be quite different (De Bandt, 1995; Kerssens-van Drongelen and Cook, 1997). To clarify how we interpret R&D performance measurement, we will first describe performance measurement as part of performance control. Secondly, we will distinguish between feedback and feed forward control, and between the R&D function and the R&D organisation. Furthermore, we will present the major design parameters of performance measurement systems and discuss the factors that have a supposed impact on these design parameters.

Performance control and performance measurement

Anthony *et al*., 1990 gave the following well-known definition of management control: 'Management control is the process by which managers influence other members of the organisation to implement the organisation's strategy.' This control process can be detailed as a set of activities, as by Kerssens-van Drongelen and Cook (1997). They formulate performance control as: '*the acquisition and analysis of information* and the *interpretation* of this information to *determine what to do and how to do it* and the *application of the chosen measures* to influence people so that their efforts are aligned to company objectives and plans.'

In accordance with this definition, *performance measurement* can be defined as the acquisition and analysis of information about the actual attainment of company objectives and plans, and about factors that may influence this attainment. In this study, we use this concept of performance measurement as part of the broader concept of performance control.

Feed forward and feedback control

Feed forward control in R&D could be interpreted as: ensuring that the right organisational conditions (qualified people, equipment, coordination mechanisms, etc.) are in place to enable good performance. Using measurement methods such as *organisational auditing* (comparing actual conditions with 'standards') or *benchmarking* against 'top-performers' or findings reported in literature, one can identify how these conditions should be changed to increase success (Cooper and Kleinschmidt, 1995; Chiesa *et al*., 1996). However, implementing these best practices does not guarantee that all R&D efforts are effective and efficient. Therefore, feedback control of performance should also be used. Feedback control can be considered as decision-making and action, based on the comparison of *objectives* with measures of *actual* performance. It should also include the comparison of these objectives and actual performance with the *expected* and the *actual internal and external conditions*. In our empirical research we focused on this performance measurement for feedback control.

The R&D function and the R&D department

When dealing with companies' R&D objectives and the subsequent measurement of performance in achieving these objectives, we find it useful to make a distinction between objectives and performance of the R&D *function* and those of the R&D *organisation*. The objective of the R&D function is to successfully initiate, coordinate and accomplish the technology process and product development activities of a company. In this definition, successfully means that the R&D function has to contribute to a company's overall (top) management function, that is: achieving those goals which determine a company's societal 'raison d'être'. As such, it is possible that the contributive function of R&D activities may be achieved without having an R&D department, e.g. through buying licenses and/or out-sourcing R&D activities. However, we realise that in practice most medium-sized and large companies do have an R&D department to support their R&D function. In this case we can formulate the objective of the R&D department as to effectively and efficiently create, sustain and exploit, at the least, the technological knowledge base needed by the company. In this formulation 'at the least' refers to the possibility of having a broader objective, e.g. the R&D department may be considered as a profit centre to perform R&D activities for other companies. The differences and overlaps of the R&D function and the R&D department are shown in Figure 1.

Responsibility and performance measurement of the R&D function. Nowadays, most R&D activities are carried out as projects (sometimes clustered within a certain research program), often under the responsibility of a *multi-functional project team* created for the duration of a specific project. Often, these teams include, besides people from the R&D organisation, staff from Marketing, Engineering, Manufacturing, Logistics, Purchasing, etc. and maybe also customer and supplier representatives. All these people together form the *R&D function* (de Weerd-Nederhof *et al.*, 1994). Within this function, the primary responsibility for a particular project *may* be in the hands of R&D. However, it is equally possible that Marketing could take the lead in new product development projects, or Manufacturing in process improvement projects.

In many companies, the *feasibility* of projects is measured periodically during the process at milestones or 'gates', for example in terms of market, strategic, economic and technical feasibility. Aspects to consider will depend upon the type of project and the project phase (van Beek, 1996; Hart *et al.* 1998). Preferably, project attractiveness is also assessed with reference to other running and potential projects in portfolio reviews (Cooper *et al.*, 1997). Furthermore, project *progress* is measured on a more frequent basis in terms of technical progress, time spent and costs, by means of a 'slip chart' or other monitoring tool (Sivathanu Pillai and Srinivasa Rao, 1996). Feasibility and progress data together can support learning and decision making about project termination or changes in project conditions or plans. These data should also be combined with data concerning *external developments and internal conditions* to assess *team performance*, that is: to judge which parts of any deviation between realisation and plan can be attributed to the team. The resulting 'team performance metrics' can serve as a basis for decision making about team rewards, changes in team constitution, etc., and moreover also as a tool to help empowered teams improve their performance and show credibility to superiors.

Responsibility and performance measurement of the R&D department. The responsibility of the R&D department overlaps the responsibility of the R&D function with respect to the effective and efficient *exploitation* of the technical competencies. However *creating* and *sustaining* the company's own technical R&D knowledge base, needed for the future, is usually not part of the responsibility of the R&D function, but only of the R&D department. On the other hand, good performance in the R&D function might also have been achieved by out-sourcing (parts of) R&D. Thus, the R&D department and each of its sub-units has to justify its existence and size by demonstrating its contribution to company performance. This calls for accurate performance measures, addressing the specific responsibilities of the R&D department. In the literature (e.g. Foster *et al.*, 1985; Moser, 1985; Brown and Svenson, 1988; Griffin and Page, 1993; Schumann *et al.* 1995; Chiesa *et al.*, 1996), we found numerous suggestions for metrics to measure R&D department

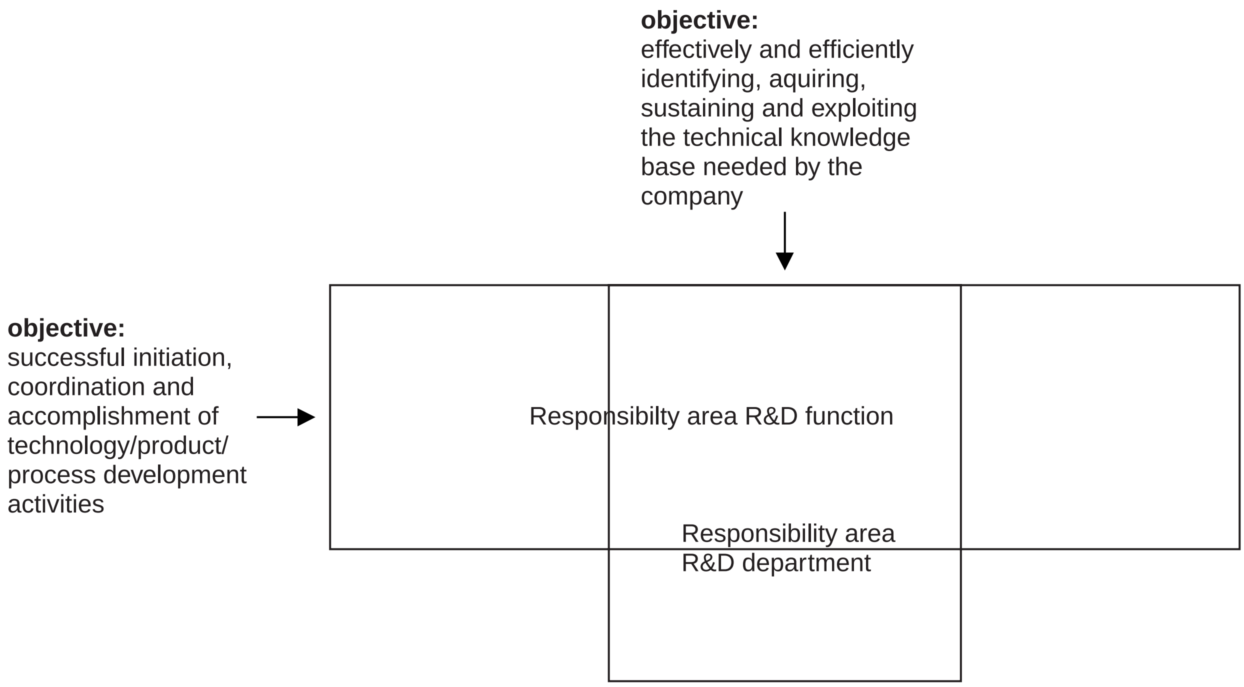

Figure 1. Differences and overlap in the responsibilities of the R&D function and the R&D department.

performance. The 'balanced scorecard' structure proposed by Kaplan and Norton (1992, 1996) seems to offer an appropriate framework to cluster these metrics, balancing measures reflecting the performance in creating and sustaining the technical knowledge base with those measures that concern the exploitation of this knowledge (see Figure 2).

The R&D department can be broken down into several sub-units: the different R&D sites, the (sub)-departments within a site, etc., down to the level of individual researchers (some levels may not exist in a specific company). From a set of metrics at the company level, subsets of lower level metrics should be derived, fitting the specific objectives and responsibilities of each sub-unit. For example, for a basic research department, traditional financial assessment may not be of use; here one would probably focus on a set of metrics derived from the internal business, and innovation and learning perspectives. The 'organisational performance measures', along with data concerning external developments and internal conditions, could assist higher level management in learning and decision making about resource allocation, reorganisation, career development and rewards etc. It could also serve as a tool for lower level empowered managers to help them improve the performance of their unit, or to demonstrate credibility to superiors and general management.

Measurement system design parameters and contingency factors

The major parameters for R&D measurement system design as described in Kerssens-van Drongelen and Cook (1997) are:

- the measures of performance (metrics) organised in a certain structure
- standards to measure performance against (norms)
- measurement techniques
- the frequency and timing of measurement and reporting
- the reporting format

The practical choices made concerning these parameters have been investigated in our survey and were discussed in further detail during the interviews.

We expected that the choices concerning these parameters would primarily be determined by the *purpose of the measurement* and the *objectives formulated for the subject of measurement*, but we also postulated that some contingency factors would have an influence. The first contingency factor considered is

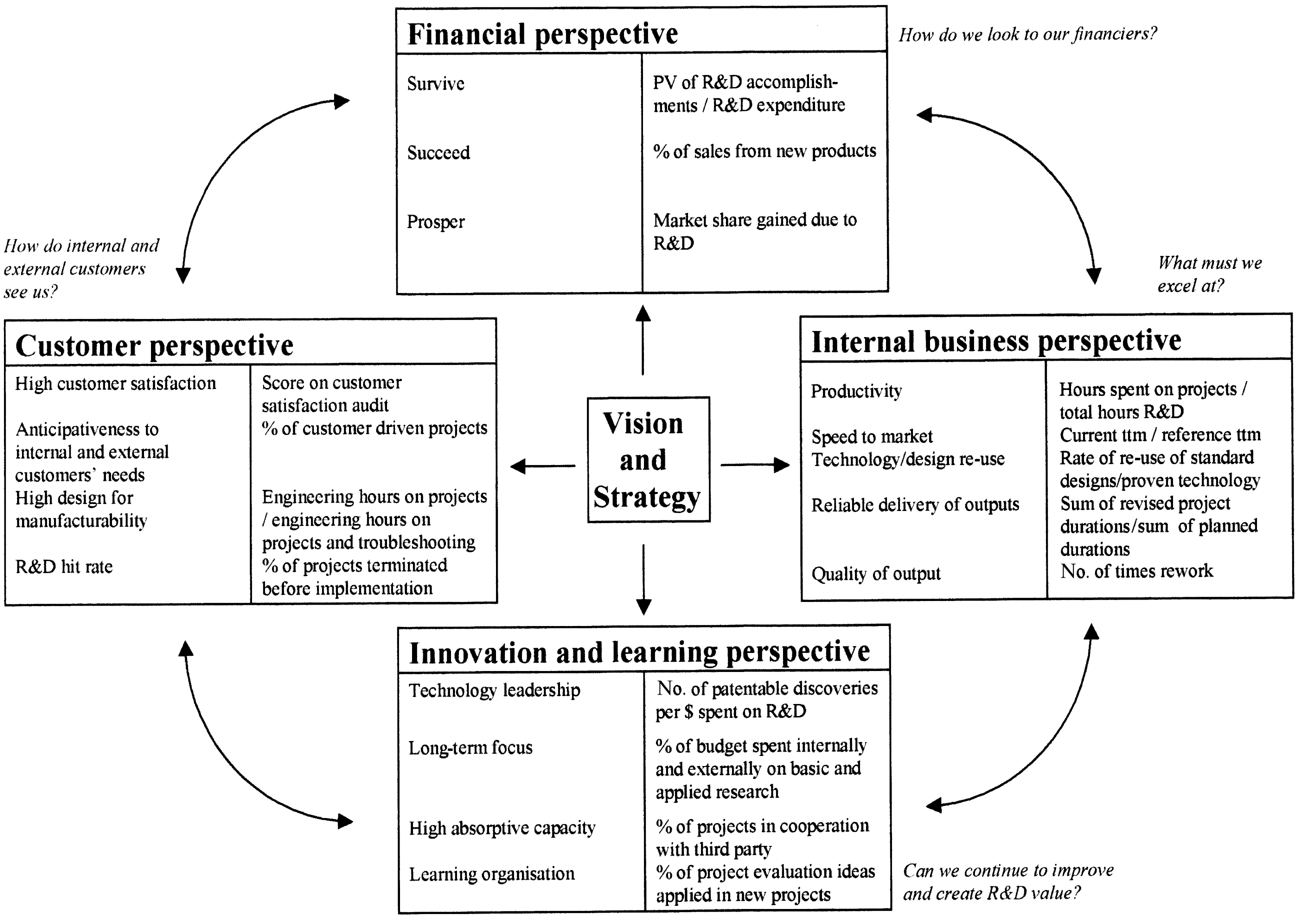

Figure 2. Example of a balanced scorecard for an R&D department
Source: Kerssens-van Drongelen and Cook, 1997.

the *organisational level* at which the measure takes place, since Ranftl (1978) found this to be of influence. Secondly, the *type of R&D*, i.e. basic research, applied research, or development, has been mentioned as having a major impact on the possibilities for measurement (Pappas and Remer, 1985; Hauser and Zettelmeyer, 1997). Thirdly, as found by Schrank *et al.* (1996), the *type of industry* influences the complexity and the organisation of R&D processes and, consequently, we postulated that this would have an impact on the measurement procedures. The fourth factor considered is *organisation size*, which is a generally acknowledged contingency factor (Emmanuel *et al.*, 1990). In the questionnaire and the interviews a distinction was made between the *company size* and the *R&D organisation size*. The final contingency factor, addressed only in the interviews, is the *strategic control model* chosen for the R&D department. Whittington (1991) gives 3 basic models for strategic R&D control:

- centralised R&D cost centre with professional (clan) control
- decentralised R&D cost centres under hierarchical business unit control
- R&D profit centres that have to acquire work from the business units or outside, thus being submitted to the market control mechanism.

These three R&D control concepts align with the general control concepts used by Ouchi (1979), who noted that each control mechanism uses different measures and measurement procedures.

Research method

As discussed earlier, the primary focus of our empirical research was on the *contingent* design of measurement systems for '*feedback measurement*' of the performance of both the *R&D function (at the individual team level)* and the *R&D department (at different organisational levels)*. This study consisted of two parts: a survey and a series of in-depth interviews with nine R&D managers. The findings from the interviews were mainly used to support, illustrate, or contrast with, the survey conclusions.

The survey sample and the selection of interviewees

As our objective for the survey was to describe rather than to explain, the sample approached was limited to 225 R&D managers. These were selected from different companies (or divisions within larger companies) in industrial sectors producing material goods, or technical wholesale businesses, known or expected to be active in R&D. 48 responses were received (21%), of which 4 companies were not active in R&D. In Table 1, the background of the remaining 44 companies is summarised. The industry breakdown in the sample follows the overall breakdown of industry sector sizes and R&D intensity in Dutch industry (Soete and Verspagen, 1993).

Of the companies involved in R&D, 80% 'measure' R&D performance in some manner. Reasons given for *not measuring* were:

- R&D activities are too integrated with manufacturing and marketing activities to be measured separately
- R&D intensity is very limited
- general management has no interest in performance measures; they assess R&D performance indirectly and subjectively
- we are still working on the design of an R&D management system

Table 1. Characteristics of the responding companies involved in R&D (n = 44).

Industry sector	Company size*	Number of employees involved in R&D (full-time equivalents)*
General chemical and chemical processing industry 25%	< 100 employees 15%	< 5 employees 18%
Pharmaceutical industry 9%	100–250 employees 30%	6–15 employees 39%
Electronics and Electro-technical industry 14%	251–500 employees 32%	16–50 employees 11%
Instruments industry 9%	501–1000 employees 7%	51–100 employees 16%
Metallurgical and machine-manufacturing industry 7%	> 1000 employees 16%	> 100 employees 11%
Transportation industry 7%		unknown 5%
Food industry 9%		
Construction and construction materials industry 9%		
Other industries 11%		

* Figures may be slightly distorted as some respondents gave only the number of employees and researchers at their own business unit, whereas others reported for the company as a whole.

- it used to be of little importance, but now we are interested in R&D measurement but do not know how to carry it out

Most of the respondents were very interested in our research: 89% wished to be informed of the results and 41% volunteered to participate in further research. From this group of volunteers, we selected nine companies for the in-depth interviews. Here we focused mainly on (subsidiaries of) larger, multi-site companies for which we assumed strategic control would be an issue. Again, a good balance between industry sectors and R&D department sizes was achieved. All interviewees were senior level R&D managers. The interviews were semi-structured and lasted between $1\frac{1}{2}$ and $2\frac{1}{2}$ hours.

Findings

Contingency factors

Measurement subject/organisational level assessed. In Table 2, a summary is given of the subjects of measurement: the R&D function (teams) and/or the R&D department. For the latter, a distinction is made between measurement at the individual level, at the level of (sub)departments, and R&D at the company level. Approximately half of the companies reported that they measured only at one organisational level, or only measured team performance. Not surprisingly these were mainly the relatively small companies in our sample. Companies with larger R&D departments mostly measure at multiple levels. Furthermore, we observed in the interviews that several companies did not perceive their individual performance evaluation system to be a type of performance measurement, whereas in our opinion it was. Thus the actual percentage of companies measuring at multiple levels is certainly higher than it first appeared.

The subject of measurement seems to have a major impact on the measurement system design, as will be illustrated in subsequent sections.

Type of R&D. The types of R&D our respondents were involved in are also summarised in Table 2. Very little basic research was done by our respondents, so our conclusions concern only applied research and development. Surprisingly, the survey findings, and the interviews, did not support our proposition on differences in measurement procedures for applied research and for development. The interviews provided a possible explanation for this in that most interviewees reported that both development and some applied research activities were performed at their site, but generally there was no departmental subdivision between development and applied research. Some companies even combined both types of R&D in one project, others preferred to have separate projects for the two types of research. But in all cases the same people carried out both types of activity. Therefore, perhaps it could be too complicated, or too confusing, to work with distinctly different measurement procedures for the two types of R&D.

Strategic control model. Among our interviewees from multi-site R&D departments we did not find any examples of centralised R&D cost centres. Two companies managed R&D as (semi-)profit centres (partly) subject to market control. This professional approach to organisational control was more or less reflected in the measurement procedures used *within* R&D: performance was measured rather objectively and compared with clear, mutually agreed, objectives. The R&D customers (BUs) were also involved in the performance assessment.

Within the category ‘decentralised R&D cost centres’ we found three variants. One company matched the classic model of decentralised R&D cost centres under hierarchical business unit control. *Within* the decentralised R&D cost centre that participated in the interview procedure, we also recorded signs of hierarchical control: detailed involvement of BU management in research planning and priority setting, ambiguous objectives and subjective assessment of performance. A variant of the decentralised cost centres model was found in two other companies: they had appointed a manager at the holding level to be responsible for preventing excessive overlap between decentralised R&D initiatives, stimulating synergy between BUs, and initiating joint R&D projects. The measurement procedures applied *within* R&D at these two companies were quite different. The third variant

Table 2. Subjects of measurement; types of research the respondents are involved in (n = 35).

Total score per subject	Score per combination of subjects	Total score per type of R&D	Score per combination of R&D activities
Team performance 54%	only team 20%	development 83%	only development 60%
Individual performance 57%	only individual 14%	applied research 40%	only applied research 14%
Department performance 26%	only total organisation 20%	basic research 6%	applied research and development 20%
Performance of R&D on company level 51%	all 4 subjects 11%		basic and applied research 3%
	various combinations 35%		all types 3%

of the decentralised cost centres model was a combination of decentralised and central cost centres. The decentralised R&D sites were primarily under local BU control, but there was also some coordination from the central R&D organisation. The decentralised R&D units with these characteristics were all re-designing their measurement procedures to be better able to assess R&D performance objectively and to empower R&D teams. Control by local BU management should become more distant as they take up the role of (internal) customer.

To summarise, we can conclude that there seems to be a link between the basic models of R&D positioning and control and the internal measurement procedures.

Other contingency factors. In Table 1 contingency factors 'type of industry', 'company size' and 'R&D department size' were listed. Unfortunately, we were not able to identify clear relationships between these contingency factors and the measurement system designs, which may be due to the limited sample size.

Purposes of measurement

As expected, the purposes of measurement clearly differed between team performance measurement and measurement at the three organisational levels (see Table 3). Table 3b shows that measurement at the sub-department and at the company level serve roughly the same purposes, albeit with different emphasis. Some overlap is also found between the purposes of team and individual performance measurement (Table 3a). Unfortunately, due to the limitations of the questionnaire design we were not able to draw clear conclusions regarding the impact of this factor on the choice of measurement system design parameters. We further observed that many respondents did not make an explicit distinction between *team performance* measurement and measurement supporting *project management*, as is shown by the high percentage of respondents stating that team performance measurement was done to control 'project progress'. This lack of distinction was also confirmed in several interviews. One interviewee complained that, due to this, R&D team performance was perceived by the company as being too low, whereas, in fact, bad project progress was caused by unfavourable and changing internal and external conditions beyond the influence of the team.

In the interviews we also discussed the link between salary increases, bonuses and other rewards, and performance measurement. At the time only one company, a subsidiary of a large American corporation, explicitly linked salary raises with individual, quantifiable performance measurements. In all other companies salary raises were based on seniority and/or more implicit links with performance assessment. All companies, except one, made occasional use of bonuses for exceptional (usually individual) performance, but in all cases this was not officially regulated. In two companies good team performance was sometimes rewarded by non-financial rewards (e.g. a party). One (also non-Dutch) company is currently developing a formal bonus system based on explicit performance measurement. However, in general the R&D managers interviewed perceived the impact on performance of intrinsic rewards and career development opportunities to be much higher than the impact of bonuses. These findings are in line with those reported by Griffin (1997). The managers' estimates of *employees*' evaluation of 'performance measurement dependent salary systems' ranged from quite positive to negative.

Table 3a. Purposes mentioned for measurement of team and individual performance.

Purposes of team performance measurement	Purposes of individual performance measurement
progress control/correction 56%	decision-making about promotion prospects 79%
decision-making to dissolve the team 28%	assignment of bonuses 47%
learning/continuous improvement 17%	correction 26%
assignment of new projects 17%	decision-making about salary 16%
assignment of bonuses 11%	decision-making about project participation 11%

Note: Numbers represent % of respondents answering the question for the indicated subject of measurement; more than one answer was possible.

Table 3b. Purposes mentioned for measurement of R&D performance at the (sub-)department and company level

Purposes of departmental performance measurement	Purposes for measurement of R&D performance at company level
assignment of resources 63%	correction 63%
decision-making about reorganisation 50%	assignment of resources 31%
correction 38%	decision-making about reorganisation 25%
assignment of new projects 25%	learning 13%
learning 13%	other 18%

Note: numbers represent % of respondents answering the question for the indicated subject of measurement; more than one answer was possible.

Measures of performance (metrics)

Unlike an earlier survey (Moser, 1985), we intentionally did not include, in the question regarding performance measures, a predetermined list of metrics. This was because we were interested to find out how the respondents themselves defined the measures, and we did not want to tempt them to list metrics they did not really use. Later, we categorised and clustered the measures mentioned along the lines of Figure 2 (see Table 4). We acknowledge that some metrics (e.g. 'number of patents') fit more than one perspective, but for reasons of simplicity we have listed them only once.

From Table 4 we have drawn the following conclusions. First, there is an almost total absence of measures from the innovation and learning perspective in team performance measurement. This would seem to support our assumption that teams are only held accountable for the proper execution of a single project, and not for long-term issues such as the generation of ideas for future business. This conclusion was confirmed during the interviews, with some interviewees believing that fostering learning, and the use of lessons learned, is explicitly a functional R&D manager's responsibility and not that of a project manager.

Secondly, it was interesting to note that two respondents, when considering the measurement of departmental performance, explicitly mentioned that the metrics used for each department differed according to the specific responsibilities. This seems to support our view that metrics ought to be aligned with the objectives and responsibilities of the measurement subject. This is also reflected in the high use score of the measure 'agreed milestones/objectives met' for team and individual performance measurement.

Thirdly, we noted that most companies did not use a balanced set of metrics from all four perspectives to measure performance at the individual, (sub)department or company level, as we suggested in our framework.

A final remark concerns the sophistication of the measures: in line with Moser (1985) we conclude that most companies indeed use broad, unsophisticated concepts such as 'quality' and 'behaviour', which are difficult to measure and interpret unambiguously.

Measurement techniques and standards

Table 5 presents a summary of the *techniques* most often used to measure the metrics, showing a balance between subjective and more objective methods. A similar balance was found in the ways performance *standards* are determined: approximately 70% of the companies use explicitly recorded standards, and subjective, implicit norms are used by 60% of the companies, especially at the individual level. The norms at this level are often based on the experiences of the evaluator, or negotiated between employee and evaluator. As could be expected, the most frequently used references, with which to compare team performance, are the terms documented in the project plan

Table 4. Measures of performance reported as used for each subject

Measures	Team	Individual	Department	Company
customer perspective:				
customer satisfaction/market response	21%	—	25%	33%
% of products succeeding in the market	5%	—	—	11%
professional esteem	5%	—	—	11%
internal business perspective:				
agreed milestones/objectives met	74%	65%	25%	—
No. projects/products completed	—	10%	63%	50%
speed	32%	50%	13%	28%
efficiency/keeping within budget	26%	10%	25%	11%
quality of output/work	42%	40%	25%	11%
behaviour (in group)	—	70%	13%	6%
planning accuracy	—	—	—	11%
innovation and learning perspective:				
No. patents	—	5%	13%	11%
No. ideas/findings	—	10%	—	6%
creativity/innovation level	5%	25%	13%	—
network building	—	5%	—	6%
financial perspective:				
expected or realised IRR/ROI	11%	—	13%	11%
% of sales by new product(s)	—	—	—	28%
profit due to R&D	—	—	—	22%
market share gained due to R&D	—	—	13%	6%

Note: numbers represent % of respondents measuring the performance of the indicated subject.

Table 5. Use of measurement techniques reported in the survey.

Measurement technique	Team	Individual	Department	Company
subjective assessment by superior(s)	32%	84%	44%	39%
assessment by independent third party	26%	5%	11%	17%
questionnaire/verbal feedback by internal and/or external customers	21%	—	44%	50%
objective score on quantitative criteria	47%	53%	67%	56%

Note: numbers represent % of respondents measuring at the indicated level; more than one answer was possible.

(time, budget, specifications). At the departmental and organisation level, a few companies based norms on a comparison with other companies, but more often past experiences or ambitions determined the standards set.

Frequency and timing of measurement

Individual performance is most often assessed on an annual basis (in 55% of all cases), mainly in a formal performance evaluation session (75%). R&D performance at the (sub)department and company level is usually measured annually (45% and 55% respectively) or monthly (about 30% in both cases). Many respondents reported that, if required, measurement can also be carried out irregularly, in particular at the departmental level (45%). Measuring R&D performance at company level is mainly carried out when the financial results (budget) are determined and reported (80%). For (sub)departments, this is also a frequently used occasion (40%), but it is more common to coincide evaluations with that of the (sub)department manager (50%). As *team performance* measurement is often considered to be the same as *project progress* measurement, it is not surprising that most respondents report that measurement takes place at milestones (45%) or project progress meetings (55%) with no fixed frequency (45%). A smaller number of the respondents measure team performance during an annual performance evaluation session (30%).

Effectiveness of the measurement procedures

Respondents' evaluation of measurement effectiveness

We acknowledge that ideally we should evaluate the value of R&D measurement procedures by their contribution to company performance. Realistically we know that the overall performance of a company is the result of numerous intertwined exogenous and endogenous factors, only one of them being the R&D process. Consequently, it seems to be impractical to isolate precisely the contribution of R&D in a quantitative and objective manner, let alone the contribution of an R&D performance measurement procedure, which is only one aspect of the organisational design of the R&D process. Therefore, a subjective perceptual measure obtained from our respondents was perceived to be the most appropriate measurement method (Hultink and Robben, 1995). Respondents were asked to evaluate their procedures in terms of the *value* they attached to them and their *impact on performance*. These two aspects were measured on a 5-point scale, 1 representing 'very high value' and 'positive impact' respectively, and 5 representing 'no value' and 'negative impact'. The average scores are indicated in Table 6.

The differences between the mean scores of the four measurement systems are barely statistically significant.[1] However, all the scores are significantly below 3 and some even below 2.5,[2] indicating that, on average, the respondents attach a medium-high value to their measurement procedures and experience a positive impact on performance. This finding supports our introductory statement that *R&D managers* no longer consider performance measurement to be inappropriate or counter productive to R&D. In the interviews we also asked for *examples* of positive impacts that the measurement procedure had had on performance. One interviewee believed that, due to the introduction of a quantitative performance measurement system (at the total R&D organisation level), researchers had become more aware of the importance of issues such as R&D productivity and reduction of development time, and were more motivated to improve performance in these areas. This has resulted in a significant improvement in R&D productivity. Another interviewee reported that, partly due to an improvement in performance measurement procedures, a reduction in development costs and a better alignment between R&D activities and business objectives have been realised.

Table 6. Evaluation of value and impact of the measurement procedures.

Measurement procedure for:	Average value attached	Average impact
team performance (n = 18)	2.1	1.8
individual performance (n = 19)	2.4	2.2
departmental performance (n = 7)	2.0	2.0
R&D performance at company level (n = 18)	2.4	2.3
average for all subjects (n = 62)	2.3	2.1

High and low rated measurement systems: what makes the difference?

In order to assess what makes the difference between measurement systems that actually improve performance and those that do not, we divided the sample into three groups. We then compared the highest rated systems (with a score of 1 for both value and impact, or scores of 1 and 2 for the two aspects) with the lowest rated systems (with scores of 4 and 3 or, at best, of 3 and 3). For *individual* performance measurement, we did not find clear differences between the best and the worst evaluated systems. However, for *team* performance measurement some interesting patterns could be observed, as shown in Table 7.

Whereas there seems to be no distinction in the metrics and norms used, the selection of the other design parameters differed. The most effective systems are characterised by measurement on a 1, 2, or 3 monthly basis, and with the involvement of representatives from Marketing or (internal) customers in the measurement procedure. The least effective team measurement systems, on the other hand, are based on measurement on a yearly or half-yearly basis, carried out solely by R&D managers. Furthermore, not only the *design parameters*, but also the *use* of the systems appears to be different. Whereas less effective systems are focused on control and correction of the evaluated project (team) only, the objectives of the most effective systems are broader, aiming also at supporting process improvement and strategic adaptation. A similar pattern was observed for *(sub)department* performance measurement systems.

At company level, we found differences as well. The highly rated systems place emphasis on metrics from the customer perspective and objective measurement techniques complemented by consultation with customers and third parties. In comparison, lowly rated systems mainly rely on metrics from the internal business perspective (in particular 'number of projects completed') and the financial perspective. The measurements are often subjective assessments by higher level managers carried out when they have to discuss results and set next year's budget. Once again, the *use* of the generated information is different. Whereas less effective performance measurement systems seem to be aimed at R&D resource allocation/budget decisions, the most effective systems are more future oriented, supporting organisational improvement processes and strategic adaptation.

Conclusions

In this paper, we have discussed the *contingent* design of measurement systems for '*feedback measurement'* of the performance of both the *R&D function (teams)* and the *R&D department* at different organisational levels. First, we explored this issue from a theoretical point of view and, as we found little empirically based literature on this subject, we carried out our own empirical study. Due to the limited sample size the conclusions of our survey are only tentative, but they are supported by in-depth interviews. Overall, the study has provided valuable material for both practitioners and academics.

Table 7. Differences between highly and hardly effective team performance measurement procedures.

	Highly effective team measurement procedure	Hardly effective team measurement procedure
System design	● measurement on 1-, 2- or 3-monthly basis ● (internal) customers involved in the measurement procedure	● measurement on a half-year or yearly basis ● measurement solely by R&D managers
Use of system output	● for correction of evaluated project (team) ● to support improvement and strategic adaptation	● for correction of evaluated project (team)

Table 8. Differences between highly and hardly effective total R&D performance measurement procedures.

	Highly effective total R&D measurement procedure	Hardly effective total R&D measurement procedure
System design	● emphasis on metrics reflecting customer demands ● objective measurement techniques and customer or third party evaluation	● emphasis on financial and output metrics (No. of projects completed) ● subjective measurement by higher level managers
Use of system output	● to support improvement and strategic adaptation	● for resource and budget allocation decisions

For R&D managers, the general overview of measurement practices in Dutch industry is a useful reference for benchmarking their own R&D measurement procedures, since they can compare themselves to the industry average. Of more interest are the differences found between effective and less effective measurement procedures for team performance measurement and R&D measurement at the company level. Frequently asking (internal) customers to evaluate R&D activities during the process, and measuring their satisfaction with R&D in general, seems to produce better performance than occasional assessments by R&D managers. These findings seem to be in line with most literature on general success factors for new product development processes, in which customer focus is high on the list (e.g. Cooper, 1996). However, in contrast with the dominant focus on metrics in the R&D performance measurement literature, our findings suggest that, at least for team performance measurement, the choice of metrics is not the determining factor. The most and least effective measurement procedures included roughly the same set of indicators. Rather, effectiveness is achieved through the combination of all the measurement system parameters listed by Kerssens-van Drongelen and Cook (1997). Furthermore, the actual usage of the measurement system output seems to be decisive. Measuring solely for diagnostic purposes does not really have an impact on performance. Performance will only improve if you actually use the information gathered in organisational improvement processes and strategic adaptation.

For academic researchers, our results offer insights into the typical level of sophistication in R&D performance measurement in large and medium-size companies. When developing new theories and frameworks for R&D performance measurement, and preparing for the implementation of these frameworks, this base level of sophistication should be kept in mind to increase the probability of adoption and success. Furthermore, we produced new evidence enabling us to draw conclusions concerning the contingent design of R&D measurement systems that are worth further investigation:

- separating the function (team) performance measurement from department performance (at different levels) appears to be appropriate
- the organisational level at which performance is measured appears to have a profound impact on the measurement procedure chosen in practice
- the strategic R&D control model adopted by the holding company appears to set the boundaries for the measurement procedures within the R&D sites

Finally, we have evidence to support our introductory proposition that performance measurement systems are perceived by R&D managers to be of value and to have a positive impact on performance.

Acknowledgements

The authors gratefully acknowledge the help of W.J. Franken, B. de Rooij, S. Slot and H.J. Veerbeek in the execution and first analysis of the survey.

References

Anthony, R.N., Dearden, J. and Govinarajan, V. (1992) *Management Control Systems: cases and readings*, 7th ed. Homewood, Illonois: Irwin.

De Bandt, J. (1995) Research and innovation: evaluation problems and procedures at different levels. *International Journal of Technology Management*, **10**, 4/5/6.

Beek, Th. A. van (1996) *Investeren in innoveren & Innoveren in investeren. Financiële Beoordeling van productinnovatie* (thesis). Technical University of Eindhoven, The Netherlands.

Brown, M.G. and Svenson, R. (1988) Measuring R&D productivity. *Research*, *Technology Management*. July–August.

Brown, S.L. and Eisenhardt, K.M. (1995) Product development: past research, present findings, and future directions. *Academy of Management Review*, **20**, 2.

Chiesa, V., Coughlan, P. and Voss, C.A. (1996) Development of a technical innovation audit. *Journal of Product Innovation Management*. **12**, 105–136.

Cooper, R.G. and Kleinschmidt, E.J. (1995) Benchmarking the firm's critical success factors in new product development. *Journal of Product Innovation Management*, **12**, 374–391.

Cooper, R.G. (1996) New products: what separates the winners from the losers'. In: Rosenau Jr, M.D. *et al.* (eds), *The PDMA Handbook of New Product Development*, New York: John Wiley & Sons.

Cooper, R.G., Edgett, S.J. and Kleinschmidt, E.J. (1997) Portfolio management in new product development: lessons from the leaders – I. *Research, Technology Management*, September-October, 16–28.

Emmanuel, C., Otley, D. and Merchant, K. (1990) *Accounting for management Control*, 2nd ed. London: Chapman and Hall.

Foster, N., Linden, L.H., Whiteley, R.L. and Kantrow, A.M. (1985) Improving the return on R&D – I'. *Research Management*, January–February.

Foster, T.M. (1996) Making R&D more effective at Westinghouse. *Research, Technology Management*, January–February, 31–37.

Francis, P.H. (1992) Putting quality into the R&D process. *Research. Technology Management*, July–August.

Griffin, A. and Page, A.L. (1993) An interim report on measuring product development success and failure. *Journal of Product Innovation Management*. **10**, 291–308.

Griffin, A. (1997) PDMA research on new product development practices: updating trends and benchmarking best practices. *Journal of Product Innovation Management*, **14**, 429–458.

Hart, S., Hultink, E.J., Tzokas, N. and Commandeur, H. (1998) Steering the new product development process: a Dutch/UK comparison. *Proceedings of the 4th*

International Workshop Meeting the Challenges of Product Development, Groningen, 16–18 April 1998.

Hauser, J.R. and Zettelmeyer, F. (1997) Metrics to evaluate R, D&E. *Research-Technology Management*, July–August, 32–38.

Hultink, E.J. and Robben, H.J. (1995) Measuring new product success: the difference that time perspective makes. *Journal of Product Innovation Management*, **12**, 392–405.

Kaplan, R.S. and Norton, D.P. (1992) The balanced scorecard – measures that drive performance. *Harvard Business* Review, January–February.

Kaplan, R.S. and Norton, D.P. (1996) Using the balanced scorecard as a strategic management system. *Harvard Business Review*, January–February.

Kerssens-van Drongelen, I.C. and Cook, A. (1997) Design principles for the development of measurement systems for research and development processes. *R&D Management*, **27**, 4, 345–357.

Kuwahara, Y. and Takeda, Y. (1990) A managerial approach to research and development cost-effectiveness evaluation. *IEEE Transactions on engineering management*, **37**, 2.

Loch, C., Stein, L. and Terwiesch, C. (1996) Measuring development performance in the electronics industry. *Journal of Product Innovation Management*, **13**, 3–20.

Moser, M.R. (1985) Measuring performance in R&D settings. *Research*, *Technology Management*, **28**, 5.

Ouchi, W.G. (1979) A conceptual framework for the design of organisational control mechanisms. *Management Science*, **25**, 9.

Pappas, R.A. and Remer, D.S. (1985) Measuring R&D productivity. *Research, Technology Management*, May–June.

Patterson, W.C. (1983) Evaluating R&D performance at the Alcoa Laboratories. *Research Management*, May–April.

Ranftl, R.M. (1978) *R&D Productivity*, 2nd ed., Hughes Aircraft Company.

Robb, W.L. (1991) How good is our research. *Research, Technology Management*, March–April.

Roussel, P.A., Saad, K.N. and Erickson, T.J. (1991) *Third Generation R&D, Managing the Link to Corporate Strategy*. Boston: Harvard Business School Press.

Schrank, R., Hollax, M. and Schug, K. (1996) Performance definition and performance measurement in R&D – results of a European–Japanese survey. *Proceedings of the R&D Management Conference on 'Quality and R&D'*, Enschede, The Netherlands, March.

Schumann, P.A. Jr, Ransley, D.L. and Prestwood, D.C.C. (1995) Measuring R&D performance. *Research, Technology Management*, May–June.

Sivathanu Pillai, A. and Srinivasa Rao, K. (1996) Performance monitoring in R&D projects. *R&D Management*, **26**, 1.

Soete, L. and Verspagen, B. (1993) De technologische positie van de Nederlandse industrie. *ESB*, **3**, 3.

Weerd-Nederhof, P.C. de, Kerssens-van Drongelen, I.C. and Verganti, R. (1994) Part I – R&D Management. *Course book 'Managing the R&D Process'*, Twente Quality Centre, Enschede/Milan.

Wheelwright, S.C. and Clark, K.B. (1992) *Revolutionizing Product Development*. New York: The Free Press.

Whittington, R. (1991) Changing control strategies in industrial R&D. *R&D Management*, **21**, 1.

Notes

1. Significance measured with *t*-tests. Only at the 10% level did the differences in impact between team performance measurement, and individual and organisational performance measurement systems become significant.
2. Measured with *t*-tests; all scores were significantly below 3 ($p < 0.01$). For team performance measurement systems, and for the total sample, the value ($p < 0.05$) and the impact ($p < 0.0005$) are both significantly below 2.5. The value score for department performance measurement ($p < 0.05$) and the impact score for individual performance measurement ($p < 0.01$) are also significantly below 2.5.