AN EXPLORATORY STUDY ON R&D PERFORMANCE MEASUREMENT PRACTICES: A SURVEY OF ITALIAN R&D-INTENSIVE FIRMS

Vittorio Chiesa, Federico Frattini, Valentina Lazzarotti, Raffaella Manzini, Ignazio Troja

1. Introduction

The role that performance measurement plays in helping firms achieve and sustain their competitive advantage has been largely acknowledged by management literature in the last decades [e.g., Eccles, 1991; Lebas, 1995; Merchant, 1998]. In the 1990s some revolutionary changes took place (i.e. technology fusion and proliferation, shortening of product life-cycles, intensified competition), that pushed top managers interest toward the measurement and assessment of R&D performance [Ortt and Smits, 2006]. Therefore R&D started to be considered accountable in terms of its efficiency, effectiveness, internal and external customer focus and alignment to corporate and business strategy [Kumpe and Bolwijn, 1994; Pearson *et al.*, 2000] and the possibility to adopt PMSs in R&D settings raised the attention of both academics and practitioners [e.g., Pappas and Remer, 1985; Brown and Svenson, 1988; Chiesa and Masella, 1996; Sivathanu and Srinivasa, 1996; Werner and Souder, 1997; Hauser, 1998; Driva and Pawar, 1999; Driva *et al.*, 2000; Poh *et al.*, 2001; Godener and Soderquist, 2004; Ojanen and Vuola, 2006]. As a result of this increased interest for the measurement of R&D activities performance, scholars started investigating the topic from different although complementary perspectives. The most relevant questions that have been addressed so far are: (i) how to design a PMS that fits the characteristics of R&D activities [Kerssens-van Drongelen and Cook, 1997; Kerssens-van Drongelen and Bilderbeek, 1999; Kerssens-van Drongelen *et al.*, 2000; Davila, 2000; Ojanen and Vuola, 2006]; (ii) how the PMS should be implemented and which are the major effects of its use [Bisbe and Otley, 2004; Godener and Soderquist, 2004]; (iii) how metrics and indicators should be selected in an uncertain and unpredictable context like R&D [Ojanen and Vuola, 2006; Chiesa *et al.* 1996; Nixon, 1998].

Empirical literature on the issue has developed as well and several case studies have been carried out on the topic.

Despite the richness of this literature, we believe that one of its major limitations is a strong bias toward the Anglo-Saxon reality, where firms are likely to be encouraged to measure the performance in their R&D units in order to satisfy accounting information disclosure requirements. To what extent and how do companies operating in countries where these type of accounting rules are absent or very weak measure the performance of their R&D activities. The paper aims at addressing this issue, focusing on a sample of Italian firms. Italy represents in fact a context where externally imposed incentives to the measurement of R&D performance are very weak. The Fourth Directive [Gray *et al.*, 1984] is the law that has regulated the disclosure of accounting information for Italian (and European) firms since the early 1980s. It does not entail that companies communicate to external stakeholders structured data about their R&D expenditures; they are merely required to provide a general description of the R&D activities that should be included in the annual report and not necessarily contain quantitative data about R&D activities, neither the simple amount of annual R&D expenses [Hall and Oriani, 2006]. In 2005, new accounting principles have been introduced in Italy, in the wake of an Europe wide harmonisation effort [Arnaboldi and Chiaroni, 2005]. These principles require that more structured data about the firm's R&D activities are disclosed [IASC, 1998], but they are far from being adopted by Italian firms, especially at the time when our empirical analysis was undertaken (first half of 2007). For these reasons, Italy represented an ideal setting for addressing the above mentioned issue.

Furthermore, the empirical literature has often failed to capture the R&D performance measurement phenomenon in its whole complexity. In other words, attempts to study from a systemic perspective the use of Performance Measurement Systems in R&D units, that accounts for the different elements the PMS is made of (e.g., objectives of the measurement, performance dimensions, control objects) are very uncommon, to the best knowledge of the authors. An exception if the work by Kerssen-van Drongelen and Bilderbeek (1999), who investigate the practical choices made by a sample of 225 Netherland firms (48 responses) concerning the major parameters for R&D measurement design. A second objective of this paper is therefore to make a step further in this direction.

On the whole, we had therefore a twofold objective: (i) to build a reference framework, through a careful review of the literature and a number of exploratory case studies, that is useful to study the problem of designing a PMS to be applied to R&D units;(ii) to study whether, why and how companies operating in contexts where institutional requirements to disclose R&D information are absent or very weak measure the performance of their R&D units. With this purpose, we used the framework developed in the first stage of the research as a reference model for a survey that involved a panel of Italian firms. We focused in particular on R&D-intensive companies, which are defined as those firms that heavily invest in and are strongly committed to R&D activities [Deeds, 2001]. We decided to focus the scope of our empirical analysis on these firms under the assumption that, because of their significant financial commitment to R&D, they will more likely feel the need to monitor these activities and hence to keep their performance under control.

The paper is structured as follows. Section 2 describes the theoretical framework that was built at the outset of our research project and used to support the empirical investigation. Section 3 develops the research questions, describes the research methodology and illustrates the details of the survey design. In Section 4 we present and discuss the results of the survey, while Section 5 concludes and outlines some directions for future research.

2. The reference framework

In this section, we synthetically develop the model we have been using as a reference framework in the survey study that is reported in this paper. In doing so, we will basically review literature adopting a systemic approach in the design of a PMS for R&D; moreover, as it is common in analytical conceptual research [Wacker, 1998], we will support our arguments using some empirical observations drawn from our exploratory case studies, that have been described in detail in previous articles [e.g., Frattini *et al.*, 2006; Chiesa *et al.*, 2007; Chiesa and Frattini, 2007].

2.1. Designing a performance measurement system for R&D units

In order to identify the constitutive elements of a PMS for R&D units, we analysed both literature on *accounting – business performance measurement* and on *R&D – innovation management*.

From the literature on accounting and business performance management [e.g., Azzone, 2006; Kaplan and Norton, 1992], we have basically drawn a *systemic* approach to performance measurement. According to this standpoint, in the design of a PMS a first fundamental step consists in clearly identifying the measurement *objectives*, i.e. the specific aims for which the system is designed and implemented. Once the objectives have been designed, it is possible to define the *dimensions of performance* to be monitored and the operative techniques to be used in order to measure the performance selected for a specified *control object.* The PMS also specifies some critical *process aspects*, basically the standards against which the measured performance should be evaluated and the frequency of the measurement. This very general reference frame can be easily applied to R&D contexts. With this purpose, we used extant literature about R&D-innovation management and performance measurement in order to operationalize the basic elements – PMS objectives, structure, dimensions of performance, process – that it comprises.

As far as the objectives of the PMS are concerned, literature has widely investigated the purposes for the use of a PMS in R&D units. The most relevant seem to be:

- Diagnosing activity for supporting resource allocation, monitoring project progress and evaluating project profitability [Kerssen-van Drongelen and Bilderbeek, 1999; Pearson *et al.*, 2000; Loch *et al.*, 1996; Bremser and Barsky, 2004];
- Motivating personnel [Kim and Oh, 2002; Kerssen van-Drongelen and Cook, 1997];
- Enhancing communication and coordination [Driva *et al.*, 2000; Loch and Tapper, 2002; Szakonyi, 1995; Bremser and Barsky, 2004];
- Learning [Driva *et al.*, 2000; Loch and Tapper, 2002];
- Reducing R&D risks and uncertainty [Chiesa and Masella, 1996; Kerssen-van Drongelen and Cook, 1997];
- Improving R&D performance [Cordero, 1990; Szakonyi, 1995].

Moreover, literature suggests that these objectives can be classified into two major categories [Azzone 2006; Chiesa *et al.*, 2007; Kerssens-van Drongelen, Cook 1997, Simons, 1994]: (i) *hard* objectives (have to ensure tight control through measurability, precision and timeliness);(ii) *soft* objectives,(have to enhance communication, information sharing and creativity). These two macro-categories of objectives lead to different choices in terms of dimensions of performance monitored by the PMS, structure of the PMS itself and process aspects. This clearly emerged in several preliminary cases that we conducted at the outset of our project [see Chiesa *et al.*, 2007].

Once measurement objectives have been defined, it is necessary to clearly specify which control objects are included in the structure of the PMS, i.e. the specific parts (e.g. units, sub-units, projects) of the firm's R&D activities and organisation whose performance are to be monitored. In R&D context, the relevant control objects identified by the literature are [Ojanen and Vuola, 2006; Kerssens-van Drongelen and Bilderbeek, 1999; Godener and Soderquist, 2004]:

- The R&D function as a whole;
- The research and the development functions separately;
- The functional units, i.e. the units responsible for R&D activities in a specific field of study or technological discipline;

- The business units' specific R&D units;
- The project teams;
- Individuals.

Obviously, these levels are not mutually exclusive; generally a PMS has a complex structure that comprises a set of interrelated control objects. Selecting proper dimensions of performance and operative indicators to be used in a PMS for R&D units is a challenging task because of the high level of uncertainty, the relevance of intangible factors, the complexity and the low standardisation of R&D activities. However, literature has deeply investigated the problem and gives us several suggestions. We focus on the main taxonomies of performance dimensions that have been advanced in the available literature. A first stream of research, particularly focused on the characteristics of R&D activities, suggests that a PMS for R&D units should comprise the following types of performance dimensions [Werner and Souder 1997; Brown and Svenson, 1988; Chiesa *et al.*, 1996; Schumann *et al.*, 1995; Sanchez and Perez, 2002]:

- Input: this means using the quantity and quality of the inputs dedicated to the operation of the control object as a proxy of its performance. Examples of input dimensions are the quantity and quality of current expenses, investments, human resources, technologies;
- Process: this means analysing the processes the control object is involved in, and perhaps responsible for, in terms of effectiveness and efficiency. Examples of typical processes that R&D control objects can be involved in are: concept generation, project selection, technology acquisition;
- Output: this means monitoring the R&D control object in terms of the actual results that it achieved, e.g. patents, scientific publications, completed projects, new products developed.

A completely different perspective is advanced by those authors who believe that the Balanced Scorecard approach could be effectively adopted in R&D units [Bremser and Barsky, 2004; Sandstrom and Toivanen, 2002; Kerssens-van Drogelen and Cook, 1997]. As a result, they suggest using the following dimensions of performance:

- Financial perspective;
- Market orientation;
- Internal business perspective;
- Innovation and learning perspective.

This is the same standpoint adopted by Neely *et al.* (2002), who study how to adapt the Performance Prism method to R&D contexts.

As far as the measurement techniques and indicators are concerned, literature has widely claimed the need to use a mix of qualitative and quantitative metrics in R&D, the former best suited to capture un-measurable aspects, and the latter capable of reducing the subjectivity of the evaluation [Pappas and Remer 1985; Werner and Souder 1997; Driva *et al.*, 2000)]. Furthermore, given that economic-financial indicators are often questionable since it is very difficult to give a monetary evaluation of uncertain and intangible elements [Frattini *et al.*, 2006], they are often integrated by non-financial indicators, whose measurement is usually easier. As far as the measurement process is concerned, a first critical issue is represented by the reference standards against which measured performance should be compared [Burch, 1994; Merchant, 1998; Kerssen van-Drongelen and Cook, 1997; Pawar and Driva, 1999; Nixon, 1998]. Again, defining standards in R&D is a really challenging issue, because of the uniqueness of each R&D project and its degree of uncertainty. In our exploratory cases, only a few companies operating in the aerospace sector were able to compare their development performance with industry benchmarks, whereas using internal past projects' performance was a relatively more diffused alternative among firms in chemical, cosmetic and pharmaceutical industries [Chiesa and Frattini, 2007]. Another critical process element is the measurement frequency. In this respect, literature underlines that it is not possible to define a single, optimal frequency [Presley and Liles, 2000; Driva *et al.*, 2000; Suomala, 2003; Pawar and Driva, 1999)], but that it should be adapted to the PMS objectives, to each dimension of performance (and/or indicator) and to different control objects. However, some general rules emerge from pieces of empirical research reported in the literature: the measurement of technical performance tends to be linked to project milestones, whilst for other dimensions a regular measurement, typically on an annual or monthly basis, is more common [Cho and Lee, 2000].

Finally, it should be stressed that the internal coherence among all the constitutive elements of the PMS is critical for ensuring its effectiveness and efficiency. This is a claim that has been widely advanced from a theoretical point of view [Azzone, 2006; Simons, 2000]. Nevertheless, literature about performance measurement in R&D units has not systematically investigated, from an empirical standpoint, the relationships among the PMS's constitutive elements. Exception are the work by Chiesa *et al.* (2007), who analysed, through an exploratory multiple case study, how the selection of the measurement objectives can affect the design of the PMS's building blocks and the paper by Kerssen-van Drongelen and Bilderbeek (1999) that we commented above. The research which is reported in this paper is believed to help fill this gap in the extant literature.

2.2. The role of contextual factors

Performance measurement research has clearly demonstrated that a close relationship exists between the PMS and the context in which performance measurement occurs.

It is possible to summarize as follows the most relevant contextual factors that should be taken into account in the design and implementation of a PMS for R&D activities, according to the available literature:

- The company's R&D strategy [Griffin and Page, 1996; Loch and Tapper, 2002];
- The type of R&D organisation [Kerssen-van Drongelen and Bilderbeek, 1999];
- The type of activities carried out (basic research and/or applied research and/or development) and the associate level of risk [Pappas and Remer, 1985];
- The resources available for the design, implementation and use of the PMS, in terms of time, money, people technology and know-how [Emmanuel et al., 1990; Godener and Soderquist, 2004];
- The company's sector of activity [Loch et al., 1996; Davila, 2000].

Table 1 summarises the most relevant relationships between the PMS's constitutive elements and contextual factors that are identified by the above mentioned literature contributions, and that we identified in our preliminary exploratory case studies [see, e.g., Frattini *et al.*, 2006; Chiesa *et al.*, 2007; Chiesa and Frattini, 2007].

PMS elements / *Contextual factors*	Objectives	Dimensions of performance	Structure	Process
R&D strategy	*	*		*
R&D organisation			*	*
Type of activity	*	*	*	*
Resources available	*	*	*	*
Sector of activity	*	*		*

Table 1: The relationships between contextual factors and PMS's elements

The theoretical development we have summarised in this Section and the exploratory cases that have been mentioned, allowed us to advance the reference framework which is reported in Figure 1.

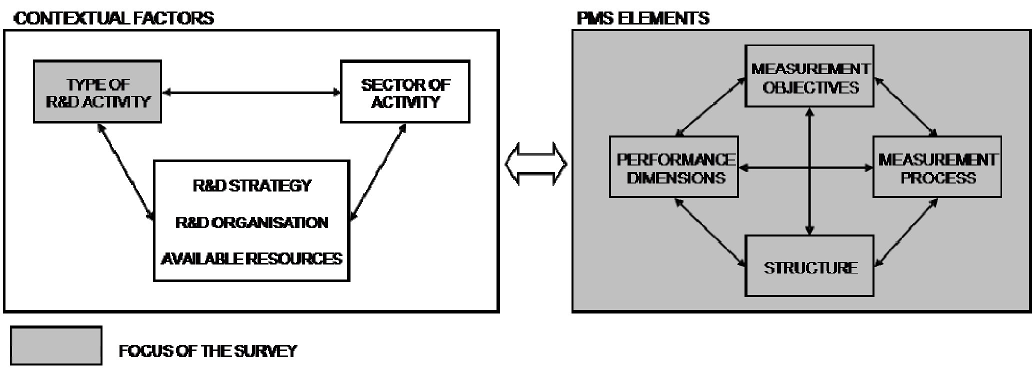

Figure 1: The reference framework and the focus of the survey

The model identifies the elements that comprise a Performance Measurement System (PMS) for R&D activities, as well as the contextual factors that influence their design. It enlightens the relationships among the PMS's constitutive elements and between constitutive elements and contextual factors.

The empirical research that is reported in this paper focuses anyhow on a limited part of this broad framework, as indicated in Figure 1. In particular, we are interested here in:

- understanding why and how Italian R&D-intensive firms measure the performance of their R&D units, i.e. how they design the constitutive elements of the PMS for R&D and which are the relationships that link these elements;
- investigating whether a relationship exists between the characteristics of the PMS for R&D designed and implemented by Italian R&D-intensive firms and the nature of the activities they undertake, being them Research or Development. The relationships between the PMS constitutive elements and the other contextual elements will be the subject of future research.

3. Research questions and methodology

In this Section we describe the fundamental research questions we wanted to answer through our empirical research, we operationalize the critical research concepts and provide methodological details about our analysis. As already mentioned, we designed and undertook a survey study, which appears to be the most adequate research methodology in light of the purpose of the paper.

3.1. Research questions

Consistently with the objectives of the paper and the theoretical framework developed in the previous Section, the survey research was designed to answer to the following questions:

- Research Question 1: Which is the degree of diffusion of Performance Measurement Systems (PMSs) in the R&D units of Italian R&D-intensive companies?;
- Research Question 2: Which are the characteristics of the PMSs designed and practically used by Italian R&D-intensive companies in their R&D units? In particular:
    - Which are the objectives for which measurement is undertaken?;
    - Which are the dimensions of performance monitored and the types of indicators used?;
    - Which is the structure of the measurement system?;

    - Which are the characteristics of the performance measurement process, both in terms of reference standards and measurement frequency?
- Research Question 3: Which are the relationships among the constitutive elements of the performance measurement systems designed and practically used by Italian R&D-intensive companies? In particular, we focus in this paper on the relationships between:
    a) Performance dimensions and all the other PMS elements, in order to understand whether a specific dimension of performance (e.g., innovation capability) is monitored with a specific typology of indicators, at a specific organisational level, against a specific type of reference standard and with a definite frequency;
    b) PMS objectives and all the other PMS elements, in order to understand whether a specific measurement objective (e.g., motivating personnel) influences the design of the associated performance dimensions, control objects, type of indicators, reference standards and measurement frequency;
- Research Question 4: Which is the relationship between the PMS and the nature of the activities undertaken in the R&D unit, namely Research and Development?. In particular, we focus here on the relationship between the objectives for which measurement is undertaken and the characteristics of the measured activities.

3.2. Operationalisation of the critical research concepts

The literature contributions that we reported in Section 2 show that each of the PMS's constitutive elements can be operationalized in a number of different ways. For instance, it is possible to cluster dimensions of performance into input, output or process measures, or to take into account the taxonomy suggested by the Balanced Scorecard approach. It is necessary, therefore, to plainly state how we operationalized the critical elements of the PMS in our empirical analysis:

- The PMS objectives were distinguished into:
    1. motivate researchers and engineers (Obj1);
    2. monitor the progress of the activities (Obj2);
    3. evaluate the profitability of projects (Obj3);
    4. select projects and investment areas (Obj4);
    5. improve R&D performance (Obj5);
    6. coordinate and communicate (Obj6);
    7. reduce uncertainty (Obj7);
    8. favor learning (Obj8).

- The dimensions of performance were classified, according to the Balanced Scorecard approach[1], into:
  1. financial perspective (P1);
  2. customer perspective (P2);
  3. efficiency of internal business processes (P3);
  4. innovation capability (P4).
- As far as the type of indicators are concerned, we distinguished into: quantitative indicators and qualitative ones
- As far as the measurement process is concerned, measurement frequencies were distinguished into: (i) periodical regular measurement; (ii) measurement by milestones. The reference standards were classified into: (i) internal reference standards, i.e. defined on the basis of the firm's past R&D performance or selected ad hoc for the specific indicator; (ii) external reference standards;
- Three possible control objects[2], which define the PMS structure, were identified: (i) the R&D unit as a whole; (ii) the R&D project; (iii) individuals.

Consistently with the research questions we wanted to answer, we designed the survey questionnaire which is reported in Appendix 1 and served as data collection tool[3].

3.3. Sample and data collection

The purpose of our research is to study the R&D performance measurement practices adopted by Italian R&D-intensive firms, which are defined as those companies that heavily invest in and are strongly committed to R&D activities (Deeds, 2001). As a population frame for our survey, we used the list of the companies that are members of the A.I.R.I. (Italian Association for Industrial Research). We eliminated from the list (which initially comprised about 170 members) those organisations that are publicly held, and we obtained a panel of 130 companies.

After a preliminary telephone contact, we sent to the R&D directors of these companies the questionnaire by mail, together with a letter describing the research project and providing a set of detailed instructions for filling in the questionnaire. We asked each R&D director to specify the type of activity ("Basic and Applied Research" or "New Product Development") undertaken in the unit he or she was responsible for (therefore, "type of activity" represented our unit of analysis). A total of 48 companies returned a properly filled in questionnaire, which resulted in 61 usable cases: 27 for Research and 34 for Development, giving a response rate of about 32%[4]. After questionnaires were returned, in order to better interpret some ambiguous answers and to ensure alternate-form reliability, we recalled about 20 firms by phone. The stability of the answers given by the whole set of respondents who were recalled in the follow up interviews indicates a good understanding of the questionnaire's items, which ensures high levels of reliability [Fink, 1995, Morgan *et al.*, 2001]

Table 2 reports the distribution of the type of activity, size[5] and industry[6] for the sampled firms. A test for non-response bias was executed using chi-square statistics, which revealed no significant differences between respondents and non-respondents[7].

		Sample	Respondents	Non-respondents
Type of activity	Research	47%	44%	**49%**
	Development	53%	56%	**51%**
Size	Small	20%	13%	**24%**
	Medium	17%	28%	**11%**
	Large	63%	59%	**65%**
Industry	Dominant design	30%	25%	**32%**
	High-tech	46%	42%	**48%**
	Science-based	**24%**	**33%**	**19%**

Table 2: Profile of the sample and the respondents

3.4. Measures

The main measures and procedures for data analysis that we applied are briefly described in the following paragraphs (see Appendix 1 for further details about the questionnaire and Appendix 2 for the values attributed to the questionnaire items). We attempted to ensure content validity by submitting the first version of our questionnaire to a combined pool of university and industry experts belonging to the "R&D Club". The experts' comments basically concerned the definition and description of the measurement objectives as well as questions' formulation and they were incorporated in the final version of the questionnaire.

3.4.1. PMS objectives

Eight objectives were defined according to the adopted conceptual framework and their relevance was measured with a 5-point scale, ranging from 1 meaning *no relevance* and 5 *very high relevance*.

3.4.2. PMS dimensions of performance, structure and process

Four performance dimensions were defined according to our framework. Their relevance was measured with a 5-point scale, with 1 meaning *no relevance* and 5 *very high relevance*. As far as the other PMS's elements (i.e. control objects, measurement frequency, typology of indicators and standards) are concerned, firms' choices were studied subordinately to each monitored performance dimension through a series of nominal variables (dichotomous or multi-categories variables).

3.4.3. Contextual factor

As far as the type of R&D activity is concerned, we applied the traditional distinction between Research and (Experimental) Development [OECD, 2002].

3.5. Methodological concerns

It should be noted that all the independent variables in our study are modelled as "attributes" of the participants but are not "active" variables, i.e. assigned to the participants themselves. As a result, we are not able to comment on the direction of the relationships among measurement objectives, performance dimensions and contextual factors that we unearthed.

In addition, the non-random nature of our sample made it impossible to apply non-parametric tests that could have been useful to verify the statistical significance of some frequencies, which we were able to interpret only in a descriptive way (see data reported in Tables 4-6b). The low frequency of many cells in the cross-tabs of Chi-square tests, moreover, often made the emerging differences between systems characterized by different objectives (i.e. evaluating profitability of projects versus motivating personnel) non-significant. Anyway, the cross-tabs' analysis allowed us to draw up Tables 7 and 8, which represent the configuration that the PMS assumed among those respondents that assigned a high relevance to the above mentioned measurement objectives.

4. Results and discussion

In this Section, we report and discuss the results of the analyses that we undertook in order to answer to each of the paper's research questions. Due to the nominal and ordinal nature of most of the variables included in the questionnaire, median and mode were the most appropriate measures of central tendency. Spearman correlations and non-parametric statistics were adopted to test the hypothesized relationships.

4.1. Research question 1

Our analysis shows that 98% of the R&D-intensive firms that responded to our questionnaire are concerned about the problem of measuring R&D performance. Although literature widely acknowledges the difficulties that practically measuring R&D performance entails, and despite the lack of external accounting requirements to disclose R&D information, our results show that R&D-intensive firms are very interested in this managerial problem and that they have been measuring their R&D performance through formalised PMSs for a number of years, often for more than a decade.

4.2. Research question 2

Table 3 displays the descriptive statistics about the measurement objectives and their achievement, together with the descriptive statistics for performance dimensions. Data indicate that half of the objectives are relevant or very relevant for the R&D-intensive firms in our sample (see mode values >= 4). This suggests that R&D-intensive firms measure the performance of their R&D with the aim to concurrently pursue several objectives. It should also be noted that the most relevant objectives (median = 5) are typically "hard": "monitor the progress of activities" and "evaluate profitability of projects". This indicates that the need for control is high even in R&D activities, that have been traditionally conceived as difficult to monitor and manage [McNair and Leibfried, 1992]. Although R&D performance measurement is essentially multi-objective, "hard" objectives appear to be more relevant and easier to achieve for R&D-intensive firms than "motivational" or "soft" ones. As far as the dimensions of performance are concerned, R&D-intensive firms generally combine multiple perspectives when it comes to measure the performance of their R&D. As Table 3 (see Appendix 3) indicates, all the perspectives suggested by the BSC approach are acknowledged to be as highly relevant (see the mode values).

Tables 4 and 5 summarise the most important results concerning the characteristics of the PMSs adopted by the R&D-intensive firms in our sample, in terms of performance dimensions, indicators, structure, reference standards and measurement frequency. The following insights are noteworthy:

- The most widely adopted control object is undoubtedly the project. Quantitative or qualitative indicators are not predominant in our analysis; this was predictable at this stage of the analysis were we consider contemporarily all the measured performance dimensions (see Section 4.3 for the design choices emerging for each specific performance dimension);

- It emerges that the performance dimensions associated to each control object are measured prevalently with a regular frequency and not in correspondence to critical project milestones. This might suggest that the use of formal stage-gate approaches for managing R&D activities, which require a set of pre-defined evaluations of a project state of advancement, is less widespread in our sample than what literature suggests [e.g., Griffin, 1997], or that performance measurement is believed to be a completely distinct practice from the stage-gate evaluations. Moreover, our results show that higher frequencies (i.e. monthly) are quite diffused only at the project level. This might suggest that firms believe the project to be the organisational object that requires tighter control than individuals or R&D unit as a whole;
- As it was largely predictable, the prevalent reference standards are internal; this confirms the difficulty in finding information about competitors' R&D performance and the lacking of industry benchmarks.

Prevalent indicators*	Prevalent Control Objects*	Prevalent frequency**	Prevalent regular frequency**	Standards*
Qualitative/Quantitative	Project	Regular	Annual	Internal

Table 4: Prevalent indicators, structure, frequency and standards (n=61). * Values calculated considering all the performance dimensions. ** Values calculated considering all the performance dimensions and all the control objects. Multiple responses were allowed.

Control object	Prevalent frequency*	Prevalent regular frequency*
R&D Unit	Regular	Annual
Project	Regular	Monthly
Individual	Regular	Annual

Table 5: Prevalent frequency for each control object (n=61). * Values calculated considering all the performance dimensions. Multiple responses were allowed.

4.3. Research question 3a

Table 6a and 6b summarize the most relevant results of our survey concerning the relationships between dimensions, indicators, control objects, frequency and standards, independently from the pursued measurement objective. Specifically, Table 6a describes the choices in terms of control objects, type of indicators and reference standards for each of the controlled performance dimension; Table 6b gives instead an overview of the measurement frequency adopted by the firms in our sample for each performance dimension and each prevalent control object as identified in Table 6a.

Some interesting insights emerge from the two Tables, even if they are obviously subject to limitations in terms of statistical significance and hence generalisability (see Section 3.5):

- The financial and market orientation perspectives are measured chiefly at the project level; on the contrary, firms assess their innovation capability mainly at the R&D unit and, interestingly, at the individual level. These results are quite obvious if we consider that financial evaluation techniques (e.g., Discounted Cash Flow approaches or real option analyses) can be properly applied only to R&D projects but not to the R&D unit as a whole, nor to the single researcher or engineer. Similarly, market orientation is a dimension of performance that can be difficultly assessed in respect to individuals or organisational units.

Performance dimension	Control object	Typology of indicator	Typology of standard
Financial perspective	Project	Quantitative	Internal
Market orientation	Project	Qualitative	Internal/External
Efficiency of R&D processes	Project	Both	Internal
Innovation capability	Individual/Unit	Qualitative	Internal/External

Table 6a: Relationships between control objects, indicators and standards for each performance dimension (n=61). Multiple responses were allowed.

- Quantitative indicators are clearly prevalent when it comes to measure the financial dimension of R&D activities, both qualitative and quantitative metrics are applied for assessing R&D efficiency, whereas qualitative ones are the most diffused with market orientation and innovation capability.
- As far as reference standards are concerned, internal ones are clearly prevalent for the financial and efficiency dimensions; although they are the most diffused also when a firm measures its R&D's innovation capability and market orientation, it is possible to note a wider use of external standards in these latter cases. In other words, respondents acknowledge that continuous and closer comparison with competitors is more critical in these instances.
- Finally, the adopted measurement frequency appears to be predominantly regular (see Table 6b); anyway, higher frequencies (specifically, monthly) are prevalent when a firms measures its R&D's financial and efficiency performance, at a project level. This finding suggests, deepening the information reported in Table 4, that firms tend to measure more closely the performance of their R&D projects along those dimensions that are more

easily quantifiable and that allow to apply structured and easily replicable evaluation methodologies.

Performance	Prevalent control object	Prevalent frequency	Prevalent regular frequency
Financial perspective	Project	Regular	Monthly
Market orientation	Project	Regular	Annual
Efficiency	Project	Regular	Monthly
Innovation capability	Individual	Regular	Annual

Table 6b: Frequency for each performance dimension/prevalent control object (as identified in Table 6a). Multiple responses were allowed.

4.4. Research questions 3b and 4

As already brought into evidence, the survey results suggest that R&D-intensive companies' PMSs are multi-objective (see Table 3 in Appendix 3); in other words, it emerges that these firms aim to achieve several different purposes by means of their R&D PMSs. This is to some extent in contrast with the available literature suggesting that *hard* and *soft* objectives correspond to significantly different choices in terms of monitored performance dimensions, structure of the PMS and measurement process aspects [Chiesa *et al.*, 2007; Azzone 2006; Kerssens-van Drongelen, Cook 1997, Simons, 1994, 2000]. In order to investigate this aspect more in detail, we undertook four types of analysis on our survey data:

a) An analysis of the correlations among different objectives, in order to verify whether some negative and/or positive relationships exist among them and, hence, whether some clusters/models of objectives can be identified (e.g., the *soft* and the *hard* clusters);
b) An analysis of the correlations among objectives and monitored performance dimensions, in order to verify whether specific objectives are linked (positively or negatively) to the dimensions of performance that are monitored;
c) An analysis of the relationships among the objectives and the type of activity (Research or Development), in order to verify whether certain objectives are pursued more intensively in a Research context rather than in a Development one, and vice versa.
d) An analysis of the elements (control objects, frequency, type of indicators and standards) that characterize the identified models;

As far as the first analysis is concerned, the following significant correlations reported in table 3 indicate that it is possible to distinguish two classes of objectives, i.e. "motivational" and "hard/diagnostic":

- the positive and significant relationship between "motivate researchers and engineers" and "improve R&D performance" (0.26, $p \leq 0.05$);

- the positive and significant relationships among “evaluate the profitability of projects”, “select projects and investment areas”, “monitor the progress of the activities” (0.4, $p \leq 0.001$; 0.24, $p \leq 0.1$);

Moreover, the negative and significant correlations between “motivate researchers and engineers” and each of the “hard” objectives (-0.21, $p \leq 0.1$; -0.30, $p \leq 0.05$; -0.32, $p \leq 0.05$), indicate that they are somehow mutually exclusive. In other words, when R&D-intensive companies try to pursue with a strong emphasis typically “hard” objectives, they generally attach a lower relevance to “motivational” ones. This can be explained considering that different objectives for performance measurement entail dissimilar measurement approaches as evidenced by the above mentioned literature. In order to understand more deeply this contrast, we analysed the results of our second analysis, which demonstrates that different objectives lead to dissimilar choices in terms of relevance of the various performance dimensions. Taking into account the correlations between measurement objectives and performance dimensions reported in Table 3, it emerges that “innovation capability” is strongly associated with “motivational” purposes (specifically, the strongest correlation is with the objective “motivate researchers” - coefficient equals 0.41, $p \leq 0.001$; a significant correlation also exists with the objective “improve R&D performance” - coefficient equals 0.30, $p \leq 0.05$). The “motivational” objective named “improve R&D performance” is positively correlated also with the “efficiency” measurement perspective (0.28, $p \leq 0.05$). In other words, it emerges that the performance of researchers and engineers is measured considering: (i) their capability to provide valuable ideas for new products or for the improvement of existing products and processes, and (ii) the efficiency with which they perform specific tasks or accomplish specific goals, e.g. the acquisition or development of new competencies. The primary need to motivate engineers and researchers brings thus R&D-intensive firms to define their accountability on the basis of those factors that deal with the influential aspects of their work and that they can completely control. This position is consistent with the theories of action, design and expectation [e.g., McClellan *et al.*, 1953; Pritchard, 1990; Moizer, 1991] which hold a relevant position in performance management literature.

On the other hand, positive and significant coefficients (0.37, $p \leq 0.01$; 0.32, $p \leq 0.05$) link the “financial” and “customer” dimensions with the “hard” objective “evaluate the profitability of project”. The “hard” objective named “monitor the progress of activities” is positively correlated also with the “efficiency” measurement perspective (0.30, $p \leq 0.05$); nevertheless, this performance dimension is probably employed in a different way than in the case it is used to “improve R&D performance”[8]. Taking into account the other “hard” objectives, only the relationship between “customer” perspective and “select projects and investment areas” shows a positive and sufficiently significant coefficient (0.24, $p \leq 0.1$). In sum, in a hard model conception, respondents assess the revenue and costs associated to R&D activities, the capacity to identify and timely incorporate into new products and processes the needs of the market, and the levels of resource consumption that they entail. In this case, in fact, the need to objectively evaluate the profitability of R&D projects, select the most profitable ones and monitor their progress, becomes of foremost importance, whereas the motivational role of performance measurement is rather neglected. Therefore, financial and quantitative indicators [Werner and Souder, 1997] and measures for key success factors such as customer satisfaction [Loch and Tapper, 2002] are introduced, although they are not completely controllable by researchers and engineers and hence are likely to have a limited motivational impact. Contrasting characteristics between the motivational model and the hard one are evidenced by the negative and significant correlation between the "financial" perspective and the objective "motivate researchers and engineers" (-0.23, $p \leq 0.1$). However, this is not true for all the couples of mutually exclusive objectives: e.g., "motivate researchers" and "select project and investment areas" are negatively and significantly correlated, but they are both positively associated with the "innovation capability" perspective.

As far as the type of activity is concerned, the Mann-Whitney U test was used to search for statistically relevant differences between the measurement objectives pursued in Research and Development units. Values of Z ($Z=-2.82$, $p \leq 0.01$) indicate a positive association between Research and the objective "motivate researchers", suggesting that in Research units motivational objectives are pursued more intensively than in Development ones. On the contrary, the "hard" objective "evaluate profitability of projects", is pursued more intensively in Development than in Research units ($Z=-2.031$, $p \leq 0.05$). No other statistically relevant differences in measurement objectives emerged. Similar results add some further suggestions about the contrast between the motivational and hard objectives. It can be explained considering that Basic and applied research is characterised by lower pressures on tangible results and deadlines than Development, by an organisational environment that favours individual work rather than teamwork, and by results that are far more distant in time. These features make it far more difficult to build and maintain a strong commitment towards the company's strategic objectives and performance measurement turns out to be a critical tool in achieving such an alignment. Moreover, pursuing hard objectives in research would be extremely difficult because of the unpredictable nature of this activity. Moreover, the scale of investments in NPD activities is much greater than in research [Cooper and Kleinschmidt, 1988], which makes the need to evaluate projects' profitability particularly compelling in development settings. Similarly, moving from research to development (and hence getting closer to market), the need to strictly monitor the performance which are critical for competition (time, costs and customer satisfaction) grows in importance.

4.5. Emerging archetypes for PMSs in R&D units

Figures 2 summarises the objectives, performance dimensions and measurement context for the "hard" and "motivational" models[9] for R&D performance measurement that emerged from the empirical analyses. R&D-intensive firms employing the so-called "hard" model pursue three major objectives (i.e. to evaluate profitability of projects, monitor the progress of activities, select project and investment areas), through a measurement system that chiefly employs the financial, customer and efficiency perspectives. This model seems to be prevalent in Development units. Companies adopting the "motivational" model aims chiefly to motivate researchers and improve R&D performance, through a measurement system that monitors innovation capability and efficiency performance dimensions, but avoid making use of the financial perspective. This archetype is relatively more diffused in Research units.

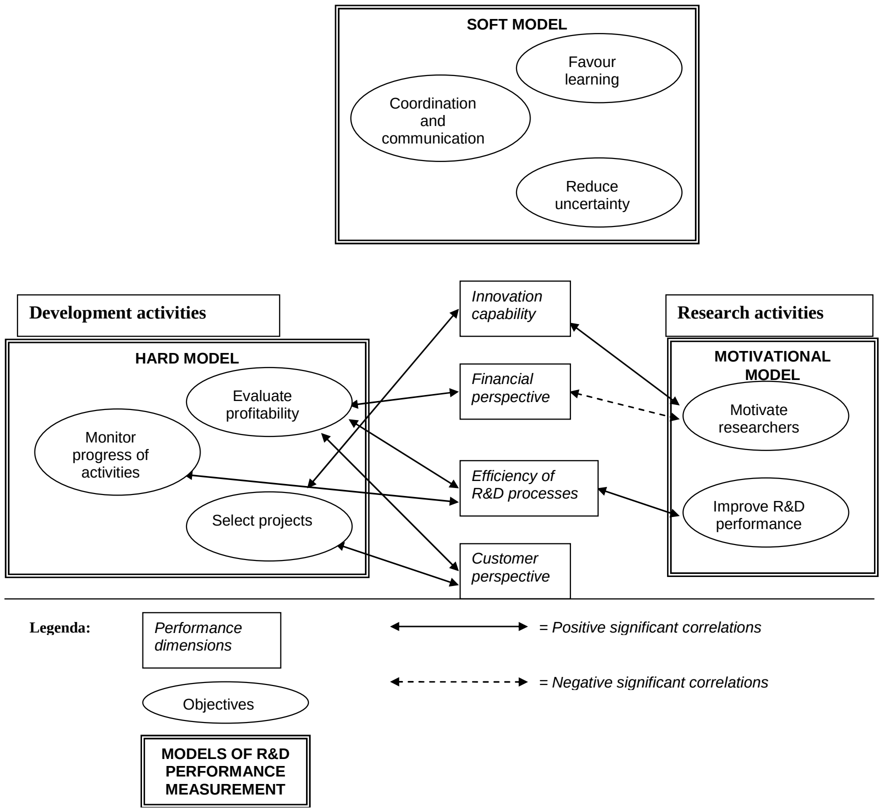

Figure 2: Objectives, performance dimensions and context for R&D performance measurement

As concerns the last analysis about the remaining PMS elements, we selected "motivational" and "hard" cases by applying the following conditions that summarise[10] the characteristics of the two models reported in Figure 2:

- *Motivational*: Obj1 (motivate researchers and engineers) ≥4 and Ob3 (evaluate the profitability of projects) <4 and P4 (innovation capability) ≥4 and P1 (financial perspective) <4;
- *Hard*: Obj3 (evaluate the profitability of projects) ≥4 and Ob1 (motivate researchers and engineers) <4 and P1 (financial perspective) ≥4 and P2 (customer perspective) ≥4.

Some interesting peculiarities of the two models emerge from the Tables 7 and 8. Even if subject to limitations in terms of statistical significance and hence generalisability (see Section 3.5), these peculiarities provide further explanations about the incompatibility of different measurement purposes:

- quite obviously, in the motivational model it is evident the prevalence of the "individual" as control object referring to its "typical" (i.e. monitored more intensively) performance dimensions (i.e. efficiency and innovation capability), while the "project" (or at most the "unit") is prevalent in the diagnostic/hard model;
- in the motivational model the frequency of the measurement is mainly annual, while the hard one shows also higher frequencies (i.e. relatively to financial and efficiency dimensions).
- motivational purposes seem to require the prevalent use of qualitative indicators at least as concerns its typical performance dimensions, while diagnostic purposes seem to be based on a wider use of quantitative information.
- no peculiarities emerge as concerns the typology of the used standards for which the general tendencies depicted in table 6a are thus confirmed.

Motivational PMS model					
Dimension of performance	**Prevalent control object**	**Prevalent frequency**	**Prevalent regular frequency**	**Prevalent indicators**	**Prevalent standards**
Financial perspective	Project	Regular	Annual	Quantitative	Internal
Market orientation	Project	Regular	Annual	Qualitative	Both
Efficiency of R&D processes	Individual	Regular	Annual	Both	Internal
Innovation capability	Individual	Regular	Annual	Qualitative	Both

Table 7: Main features of the "motivational" PMS model

Hard PMS model					
Dimension of performance	**Prevalent control object**	**Prevalent frequency**	**Prevalent regular frequency**	**Prevalent indicators**	**Prevalent standards**
Financial perspective	Project	Regular	Monthly	Quantitative	Internal
Market orientation	Project	Regular	Annual	Qualitative	Both
Efficiency of R&D processes	Project	Regular	Monthly	Quantitative	Internal
Innovation capability	Unit	Regular	Annual	Both	Both

Table 8: Main features of the “hard” PMS model.

5. Conclusions

The paper reports the results of a survey aimed at investigating the use and the characteristics of Performance Measurement Systems (PMSs) applied in the R&D units by a panel of Italian R&D-intensive firms. In light of the weakness of the R&D information disclosure requirements Italian companies are subject to, the results of our empirical analysis suggest how R&D-intensive firms design and use PMSs in R&D with the purpose to support their management activities.

It is believed that our paper contributes to the extant research on performance measurement in R&D in a number of ways. First of all, it provides a systematic framework that identifies the logical steps a firm should go through when designing a PMS to be applied to its R&D units. It is believed that this framework could inform future theoretical and empirical research into the matter, e.g., the investigation of the effects that further contextual variables (besides the type of R&D activities undertaken) have on the design of the PMS’s constitutive elements. Furthermore, the paper represents one of the first contributions, to the best knowledge of the authors, that investigate the diffusion and the characteristics of PMSs applied in the R&D units of Italian firms and, more broadly, in a country where R&D disclosure information requirements are absent or very weak. In this respect, our findings might be representative of other countries where this type of normative environment is present (e.g., Germany, among the European countries). Further longitudinal studies of R&D performance measurement practices in other countries outside Italy are required to externally validate our results.

Besides the exploratory nature of our research, it is believed that it holds some interesting managerial implications as well. First of all, the reference framework that is developed can guide R&D managers in their PMS design efforts, indicating which are the aspects and contextual factors they should focus on when designing a measurement system for their R&D units. Furthermore, the results of our empirical analysis provide a number of insights about how R&D-intensive firms operatively measure the performance in their R&D units for managerial support purposes. In particular, the emergence of the two paradigmatic models for the R&D PMS and the prevalence of these models in Research and Development settings is especially insightful in this respect. In order to improve the managerial usefulness of this paper, it is necessary to evaluate the effectiveness of the different paradigmatic models, in different measurement contexts. This represents a further promising avenue for future research.

Bibliografia

Arnaboldi, M. and Chiaroni, D. (2006) I nuovi standard contabili internazionali IAS-IFRS. Principi e casi reali. Milano: Polipress.

Azzone, G.(2006) Sistemi di controllo di gestione. Metodi, strumenti e applicazioni. Milano: Etas.

Bisbe, J. and Otley, D. (2004) The effects of the interactive use of management control systems on product innovation. Accounting, Organizations and Society, 29, 8, 709-737.

Bititci, U.S. (2000) Dynamics of performance measurement systems. International Journal of Operations Management, 20, 6, 692-704.

Bremser, W.G. and Barsky, N.P. (2004) Utilizing the Balanced Scorecard for R&D Performance Measurement. R&D Management, 34, 3, 229-238.

Brown, M.G. and Svenson, R.A. (1988) Measuring R&D Productivity. Research-Technology Management, 31, 4, 11-15.

Brown, M.G. and Svenson, R.A. (1998) Measuring R&D Productivity. Research-Technology Management, 41, 6, 30-35.

Burch, J. (1994) Cost and Management Accounting - A Modern Approach. Saint Paul: West.

Chesbrough, H. (2003) Open Innovation: The New Imperative for Creating and Profiting from Technology, Boston: Harvard Business School Press.

Chiesa, V., Coughlan, P. and Voss, C. (1996) Development of a technical innovation audit. Journal of Product Innovation Management, 13, 2, 105-136.

Chiesa, V. and Masella, C. (1996) Searching for an effective measure of R&D performance. Management Decision, 34, 7, 49-57.

Chiesa, V. and Frattini, F. (2007) Exploring the differences in performance measurement between research and development: evidence from a multiple case study. R&D Management, 37, 283-301.

Chiesa, V., Frattini, F., Lazzarotti, V. and Manzini, R. (2006) How to measure R&D performance: a design framework and an empirical study, in Manzoni, J.F. and Epstein, M.J. (eds.), Performance Measurement and Management Control: Improving Organizations and Society. Amsterdam: Elsevier.

Chiesa, V., Frattini, F., Lazzarotti, V. and Manzini, R. (2007) How do measurement objectives influence the R&D performance measurement system design? Evidence from a multiple case study. Management Research News, 30, 3, 187-202.

Cho, E. and Lee, M. (2005) An exploratory study on contingency factors affecting R&D performance measurement. International Journal of Manpower, 26, 6, 502-512.

Cooper, R. and Kleinschmidt, E. (1988) Resource allocation in the new product process. Industrial Marketing Management, 17, 259-262.

Cordero, R., (1990) The measurement of innovation performance in the firm: an overview. Research Policy, 19, 185-192.

Davila, T. (2000) An empirical study on the drivers of management control systems' design in new product development. Accounting, Organizations and Society, 25, 383-409.

Deeds, D.L. (2001) The role of R&D intensity, technical development and absorptive capacity in creating entrepreneurial wealth in high-tech start-up. Journal of Engineering and Technology Management, 18, 29-47.

Driva, H. and Pawar, K.S. (1999) Performance measurement for product design and development in a manufacturing environment. International Journal of Production Economics, 60, 61-68.

Driva, H., Pawar, K.S. and Menon, U. (2000) Measuring product development performance in manufacturing organisations. International Journal of Production Economics, 63, 2, 147-159.

Eccles, R.G. (1991) Performance measurement manifesto. Harvard Business Review, 69, 131-137.

Eccles, R.G. and Pyburn, P.J. (1992) Creating a comprehensive system to measure performance. Management Accounting, 74, 41-44.

Egelhoff, W.G. (1982) Strategy and structure in multinational corporations: an information-processing approach. Administrative Science Quarterly, 27, 435-458.

Emmanuel, C., Otley, D. and Merchant, K. (1990) Accounting for Management Control. London: Chapman & Hall.

OECD (2002) Modifications to Recommendation 96/280/CE with reference to definition of small and medium firms. Bruxelles, June, 25.

Fink, A.(1995). How to analyze survey data. Thousand Oaks: Sage.

Frattini, F., Lazzarotti, V. and Manzini, R. (2006) Towards a performance measurement system for the research activities: NiKem Research case study. International Journal of Innovation Management, 10, 4, 425-454.

Godener, A. and Soderquist, K.E. (2004) Use and impact of performance measurement results in R&D and NPD: an exploratory study. R&D Management, 34, 191-219.

Gray, S.J., and Coenenberg, A. G. (eds.) EEC accounting and harmonisation: implementation and impact of the fourth directive. Amsterdam: North-Holland – Elsevier Science Publishers.

Gregory, M.J. (1993) Integrated performance measurement. International Journal of Production Economics, 30/31, 281-296.

Griffin, A. and Page, A. (1993) An Interim Report on Measuring Product Development Success and Failure. Journal of Product Innovation Management, 10, 291-308.

Griffin, A. and Page, A. (1996) PDMA Success Measurement Project: Recommended Measures for Product Development Success and Failure. Journal of Product Innovation Management, 13, 478-496.

Griffin, A. (1997) Drivers of NPD success: the 1997 PDMA report. Chicago: Product Development and Management Association.

Hall, B.H. and Oriani, R. (2006) Does the market value of R&D investments by European firms? Evidence from a panel of manufacturing firms in France, Germany and Italy. International Journal of Industrial Organisation, 24, 971-993.

Hauser, J.R. (1998) Research, development and engineering metrics. Management Science, 44 12, 1670-1689.

IASC (1998) IAS 38: Intangible Assets.

Kaplan R.S. and Norton, D.P. (1992) The Balanced Scorecard – measures that drive performance. Harvard Business Review, 1, 171-179.

Kerssen-van Drongelen, I.C. and Bilderbeek, J. (1999) R&D performance measurement: more than choosing a set of metrics. R&D Management, 29, 1, 35-46.

Kerssen-van Drongelen I.C., Nixon, B. and Pearson, A. (2000) Performance measurement in industrial R&D. International Journal of Management Reviews, 2, 111-143.

Kerssen-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.

Kim, B. and Oh, H. (2002) An effective R&D PMS: survey of Korean R&D researchers. International Journal of Management Science, 30, 19-31.

Kodama, F. (1991) Emerging Patterns of Innovation. Sources of Japan's Technological Edge. Boston, Massachusetts: Harvard Business School Press.

Kumpe, T. and Bolwijn, P.T. (1994) Towards the innovative firm-challenge for R&D management. Research-Technology Management, 37, 1, 38-44.

Lebas, M.J. (1995) Performance measurement and performance management. International Journal of Production economics, 41, 23-35.

Legare, T.L. (2001) How Hewlett-Packard used virtual cross-functional teams to deliver healthcare industry solutions. Journal of Organizational Excellence, 20, 4, 29-38.

Loch, C.H. and Tapper, S. (2002) Implementing a strategy-driven performance measurement system for an applied research group. Journal of Product Innovation Management, 19, 3, 185-198.

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

Lutz, R.A. (1994) Implementing technological change with cross-functional teams. Research-Technology Management, 37, 2, 14-18.

McClellan, D.C., Atkinson, J., Clark, R.A. and Lowell, E.L. (1953) The achievement motive. New York: Appleton century Crofts.

McNair, C.J. and Leibfried, K.H.J. (1992). Benchmarking: a tool for continuous improvement, New York: Harper Business.

Meyer, M.H., Tertzakian, P. and Utterback, J.M. (1997) Metrics for Managing Research and Development in the Context of the Product family. Management Science, 43, 1, 88-111.

Merchant, K.A. (1998) Modern Management Control Systems: Text and Cases. Englewood Cliffs: Prentice Hall.

Morgan, G.A., Griego, O.V. and Gloeckner, G.W. (2001). SPSS For Windows: An Introduction to Use and Interpretation in Research. London: Lawrence Erlbaum Associates Publishers.

Moizer, P. (1991) Performance appraisal and rewards, in Ashton, D., Hopper, T. and Scapens, R.W., (eds.), Issues in Management Accounting, Englewood Cliffs: Prentice Hall, 126-144.

Nayak, P.R., 1987. Measuring product creation effectiveness. Journal of Business Strategy, 13, 48-52.

Neely, A., Gregory, M. and Platts, K. (1995) A performance measurement systems design: a literature review and research agenda. International Journal of Operations and Production Management, 15, 4, 80-116.

Neely, A., Bourne, M., Platts, K. and Mills, J. (2002) The success and failure of performance measurement initiatives, Perceptions of participating managers. International Journal of Operations & Production Management, 22, 1288-1310.

Nixon, B. (1998) Research and development performance measurement: a case study. Management Accounting Research, 9, 329-355.

Ojanen V. and Vuola O. (2006) Coping with the multiple dimensions of R&D performance analysis.. International Journal of Technology Management, 33, 279-290.

Ojanen, V. and Tuominen, M. (2002) An analytic approach to measuring the overall effectiveness of R&D–a case study in the telecom sector. Proceedings: Volume II of IEMC 2002, International Engineering Management Conference, Cambridge, UK, 667–672.

Ortt, J.R. and Smits, R. (2006) Innovation management: different approaches to cope with the same trends. International Journal of Technology Management, 34, 3/4, 296-318.

Pappas, R.A. and Remer, D.S. (1985) Measuring R&D productivity. Research-Technology Management, 28, 3, 15-22.

Pawar, K.S. and Driva, H. (1999) Performance measurement for product design and development in a manufacturing environment. International Journal of Production Economics, 1, 61-68.

Pearson, A.W., Nixon, W.A. and Kerssen-van Drongelen, I.C. (2000) R&D as a business - what are the implications for performance measurement?. R&D Management, 30, 4, 355-366.

Pillai. A.S., Joshi, A., Rao, K.S., 2001. Performance measurement of R&D projects in a multi-project, concurrent engineering environment. International Journal of Project Management, 20, 165-177.

Poh, K.L., Ang, B.W. and Bai, F. (2001) A comparative analysis of R&D project evaluation methods. R&D Management, 31, 1, 63-76.

Presley, A. and Liles, D. (2000) R&D validation planning: a methodology to link technical validations to benefits measurement. R&D Management, 30, 1, 55-65.

Pritchard, R.D. (1990) Measuring and Improving Organizational Productivity: a Practical Guide New York: Praeger.

Sanchez, A.M. and Perez, M.P. (2002) R&D project efficiency management in the Spanish industry. International Journal of Project Management, 20, 7, 545-560.

Sandstrom, J. and Toivanen, J. (2002) The problem of managing product development engineers: Can the balanced scorecard be an answer?. International Journal of Production Economics, 78, 79-90.

Schumann, P.A., Ransley, D.L. and Prestwood, D.C.L. (1995) Measuring R&D Performance. Research-Technology Management, 38, 3, 45-54.

Shenhar, A. and Dvir, D. (1996) Towards a typological theory of project management. Research Policy, 25, 607-632.

Simons, R. (1994) Levers of control: how managers use innovative control systems to drive strategic renewal. Boston: Harvard Business School Press.

Simons, R. (2000) Performance Measurement & Control Systems for Implementing Strategy. Prentice Hall.

Sivathanu, P.A. and Srinivasa, R.K. (1996) Performance monitoring in R&D projects. R&D Management, 26, 1, 57-65.

Suomala, P. (2003) Multifaceted New Product Development Performance: Survey on Utilization of Performance Measures in Finnish Industry. Proceedings of The 2nd Workshop on Performance Measurement and Management Control, Nice, France. 18.-19 September.

Szakonyi, R. (1995) Measuring R&D effectiveness-I. Research-Technology Management, 37, 27-32.

Tushman, M.L. and Nadler, D.A. (1978) Information-processing as an integrating concept in organizational design. Academy of Management Review, 3, 3, 613-624.

Valle, S. and Avella, L. (2003) Cross-functionality and leadership of the new product development teams. European Journal of Innovation Management, 6, 1, 32-47.

Wacker, J.G. (1998) A definition of theory: research guidelines for different theory-building research methods in operations management Journal of Operations Management, 16, 4, 361-385.

Werner, B.M. and Souder, W.E. (1997) Measuring R&D performance - State of the art. Research-

Appendix 1.

The questionnaire

Sector of activity	
Type of activity (Research or development[11])	
Turnover	
Number of employees	

1. Are you interested in the measurement of R&D performance?

   ☐ Yes
   ☐ No

2. If yes, do you measure R&D performance in a systematic and formalised way?

   ☐ Yes
   ☐ No

3. If yes, how long has been measured R&D performance?

   ☐ Less than five years
   ☐ Five to ten years
   ☐ More than ten years

4. Why do you measure? (indicate the relevance of the each objective: 1=none; 5=very high)

Motivating personnel	1	2	3	4	5
Monitoring activities in progress (time and costs)	1	2	3	4	5
Valuating projects' profitability	1	2	3	4	5
Selecting investment projects and areas	1	2	3	4	5
Improving R&D effectiveness	1	2	3	4	5
Improving communication and coordination	1	2	3	4	5
Reducing uncertainty[12]	1	2	3	4	5
Stimulating learning[13]	1	2	3	4	5

5. Which dimensions of performance do you actually measure? (Indicate the relevance of each dimension: 1=none; 5=very high)

Financial performance	1	2	3	4	5
Market Orientation	1	2	3	4	5
Efficiency of the R&D processes	1	2	3	4	5
Innovation capability	1	2	3	4	5

6. Which are the control objects of the measurement, i.e., which is the organisational level of the measurement? U = R&D Unit; P = Project; I = Individual (you can choose more than a single control object for each dimension of performance)

Financial performance	U	P	I
Market Orientation	U	P	I
Efficiency of the R&D processes	U	P	I
Innovation capability	U	P	I

7. Which is the frequency of the measurement? Regular (1= weekly; 2= monthly; 3= quarterly; 4=six months; 5= yearly) or by milestones? (you can choose more than a single answer for each dimension of performance)

		Regular					Milestones
Financial perspective	**U**	1	2	3	4	5	☐
	P	1	2	3	4	5	☐
	I	1	2	3	4	5	☐
Market orientation	**U**	1	2	3	4	5	☐
	P	1	2	3	4	5	☐
	I	1	2	3	4	5	☐
Efficiency of the R&D processes	**U**	1	2	3	4	5	☐
	P	1	2	3	4	5	☐
	I	1	2	3	4	5	☐
Innovation capability	**U**	1	2	3	4	5	☐
	P	1	2	3	4	5	☐
	I	1	2	3	4	5	☐
	P	1	2	3	4	5	☐
	I	1	2	3	4	5	☐

8. Which kind of indicators do you use to measure the performance? QT= Quantitative[14]; QL = Qualitative[15] (you can choose both types of indicator for each dimension of performance)

Financial performance	QT	QL
Market Orientation	QT	QL
Efficiency of the R&D processes	QT	QL
Innovation capability	QT	QL

9. Which kind of reference standard do you use to evaluate the performance measured? RI= Internal reference[16;] RE= External reference[17] (you can choose both types of standard for each dimension of performance)

Financial performance	RI	RE
Market Orientation	RI	RE
Efficiency of the R&D processes	RI	RE
Innovation capability	RI	RE

10. Do you think that the objectives that you have identified for your performance measurement have been actually achieved? (1= at a very low level; 5= at a very high level)

Motivating personnel	1	2	3	4	5
Monitoring activities in progress (time and costs)	1	2	3	4	5
Valuating projects' profitability	1	2	3	4	5
Selecting investment projects and areas	1	2	3	4	5
Improving R&D effectiveness	1	2	3	4	5
Improving communication and coordination	1	2	3	4	5
Reducing uncertainty[18]	1	2	3	4	5
Stimulating learning[19]	1	2	3	4	5

Examples of indicators for each dimension of performance

Performance dimensions	**Indicator**
Financial performance	• Total cost of each project • R&D annual spending • R&D annual investment • R&D annual costs • PV of R&D accomplishments / R&D expenditure • R&D or technology acquisition cost per new product or project • IRR or NPV due to R&D projects • ROI due to R&D projects • Sales (or % of sales) by new products • Cost (or % of cost) reduction by new projects • Profits due to R&D • Market share due to R&D and innovations • Cost/benefits performance of completed R&D projects
Market orientation	• Promotional spending • % of customer driven projects • Accuracy of prediction and interpretation of customer requirements (measured by: n. of interaction with customer during the project; n. of customers included in the project team; % budget dedicated • to customer analysis or verification) • Time to market • N. of presentation to external customers • N. of training sessions signed off by customer and delivered • N. of problem analysis reports requested and delivered • Engineering hours on projects / engineering hours on projects and troubleshooting • N. of new customers • N. of customer complaints • Customer satisfaction (n. or % of design meeting customer needs; product range and variety) • % of products succeeding in the market (success intended as the achievement of sales objectives) • % of support requests fulfilled • Response time to customer requests for "specials"
Efficiency of R&D processes	• Experience of R&D employees • Hours or % budget • N. or % of employees involved in goal setting • Availability (knowledge) of advanced managerial tools (project management techniques • Availability (knowledge) of advanced IT support tools (e.g., rapid prototyping tools, design support tools etc.) • Average annual improvement in process parameters (quality cost, lead time, wip, reliability, capability, down time) • N. or % of certified processes • % of projects that lead to new or enhanced products or processes • % of R&D expenditure that lead to new or enhanced products or processes • % Design using techniques as design for assembly, design for manufacturing, design to cost ….) • Rate of re-use of standard designs / proven technology • Sum of revised project durations / sum of planned durations • Quality of documentation to development (measured by, for example, % phases / process documented) • Agreed milestones / objectives met • Time spent on changes to original product/project specification • % of projects using a common design platform • % of projects respecting costs and budget • % of projects delayed or cancelled due to lack of funding • % of projects delayed or cancelled due to lack of human resources • Installation lead time • Average overrun

	• Average time of re-design • Average time of product enhancement • N. of times rework • N. or % of products / projects completed • Rate of successful projects (i.e. project achieving the assigned time, cost, quality) • % of project milestones completed • Total product development time • Product quality (measured through indicators specific for each industry / product) • % on time delivery of specification to manufacturing • N. of customer detected design faults
Innovation capability	• N. or % people with management experience (i.e. adequate degree or dedicated training) • N. of employees in R&D • % of suggestions implemented • N. of meeting or time dedicated to the analysis of reasons for failure of previous projects • N. of hours of staff training • N. of products in development or projects in course • N. of publications • N. of scientific awards • N. of patents registered/pending • N. of new ideas per year • N. of innovations delivered to production and commercialization • N. of requests handbooks published and delivered • N. of new product-based business area/ventures started in past 5 years • Average product life-cycle length • N. of new processes and significant enhancements per year • N. of improvements suggestions per employee

Appendix 2: Values assigned to the main questionnaire items

	Value	Label
Interest	0	No
	1	Yes
System formalization	0	No
	1	Yes
Age	1	Less than 5
	2	5-10
	3	More than 10
Score to assess objectives' relevance	1	No relevance
	2	Low relevance
	3	Medium relevance
	4	High relevance
	5	Very high relevance
Score to assess performance dimensions' relevance	1	No relevance
	2	Low relevance
	3	Medium relevance
	4	High relevance
	5	Very high relevance
Frequency	0	Milestones
	1	Regular
	2	Both
Regular frequency	1	Weekly
	2	Monthly
	3	Quarterly
	4	Semi-annual
	5	Annual
Indicators	1	Quantitative
	2	Qualitative
	3	Both
Standards	1	Internal
	2	External
	3	Both
Score to assess objectives' success	1	Very low
	2	Low
	3	Medium
	4	High
	5	Very high

Appendix 3: Table 3

Variables	Median	Interquartile Range	Mode	Obj1	Obj2	Obj3	Obj4	Obj5	Obj6	Obj7	Obj8	p1	p2	p3	p4
Measurement purposes:															
Obj1	3	2	3	1	**-0.21×**	**-0.30***	**-0.32***	**0.26***	0.01	0.10	**0.21×**	**-0.23×**	0.06	0.08	**0.41*****
Obj2	5	1	5	**-0.21×**	1	**0.40*****	0.19	0.07	**0.23×**	0.20	**0.30***	0.05	0.15	**0.30***	**0.11**
Obj3	5	2	5	**-0.30***	**0.40*****	1	**0.24×**	0.11	0.12	0.10	0.09	**0.37****	**0.32***	**0.21×**	**0.02**
Obj4	4	1.5	4	**-0.32***	0.19	**0.24×**	1	0.07	0.03	0.07	0.12	0.16	**0.24×**	0.12	**0.30***
Obj5	4	1	4	**0.26***	0.07	0.11	0.07	1	0.10	0.10	**0.22×**	-0.02	0.08	**0.28***	**0.30***
Obj6	3	1	3	0.01	**0.23×**	0.12	0.03	0.10	1	**0.31***	**0.54*****	0.10	-0.03	**0.30***	**0.06**
Obj7	3	2	3	0.10	0.20	0.10	0.07	0.10	**0.31***	1	**0.37****	0.20	0.08	0.07	**0.21**
Obj8	3	2	3	**0.21×**	**0.30***	0.09	0.12	**0.22×**	**0.54*****	**0.37****	1	-0.05	0.12	**0.35****	**0.25×**
Accomplishment:															
Obj1	3	1	3												
Obj2	4	1	4												
Obj3	4	1	4												
Obj4	4	1	4												
Obj5	4	1	4												
Obj6	3	1	3												
Obj7	3	1	3												
Obj8	3	2	3												
Performance:															
P1	4	1	4												
P2	4	1	4												
P3	4	2	4												
P4	4	2	4												

Notes

1 We decided to apply the BSC approach because literature has already shown that it can be effectively employed as a reference framework to assess the performance of R&D units [Bremser and Barsky, 2004]. It should be noted that we provided respondents, in a document attached to the questionnaire, with a list of metrics that were representative of the different performance dimensions included in the BSC approach. These metrics were developed through a literature review and were discussed with the components of the "R&D Club", an interest group that gathers researchers and professors from 4 Italian Universities and R&D managers from a selection of the leading Italian firms. They meet twice a year with the purpose to discuss relevant issues in the field of R&D management. Furthermore, the 20 follow-up telephone interviews conducted after the questionnaires were returned (as discussed in Section 3.3) confirmed that our respondents had properly understood the meaning of the questionnaire's items

2 The decision to choose these three control objects to be included in our analysis, among those suggested by the literature, is motivated by the unit of analysis adopted in our research, which is the PMS applied to the R&D units of Italian R&D-intensive companies. We are not interested therefore in higher-level (e.g., organisational) measurement approaches, neither to those that take into account the relationships the relevant firm establishes with external organisations with which it collaborates in the wake of the well-known Open Innovation paradigm [Chesbrough, 2003]. In fact, we consider the ability of the R&D unit to collaborate with external organisations through the innovation capability performance dimension, and especially the indicators we used to indicate to our respondents what we meant exactly with this performance dimensions.

3 As already mentioned speaking about the choice of the BSC approach, in order to improve the reliability of the questionnaire, we accompanied each question about the importance of measurement objectives and performance dimensions with a set of real world examples or paradigmatic situations, that were specifically developed to illustrate the meaning that we attached to each item. These examples were developed together with the experts taking part to the "R&D Club" and helped enhance the reliability of the research, that was confirmed by the 20 follow up telephone interviews we mentioned in Section 3.3.

4 13 firms in our respondents had both Research and Development activities. Considering the available information about the "type of activity" declared by companies, the 130 companies included in the sample resulted in a total of 191 cases.

5 We applied the criteria suggested by the EU (European Commission Recommendation, 2002) for classifying firms on the basis of their size. Specifically, a company was classified as: (i) small, if the number of workers is < 50 and the revenues < 10 million €; (ii) medium, if the number of workers is between 50 and 250 and the revenues 10-50 million €; (iii) large, if the number of workers is > 250 and the revenues > 50 million €.

6 Industrial sectors were classified, adopting the taxonomy proposed by Kodama (1991), into Dominant Design, High Tech and Science Based industries. These three clusters differ in respect to their level of risk, defined as "the probability that an R&D program expenditure is frozen" in an intermediate phase of the R&D process (freezing rate). Dominant design industries are characterised by a freezing rate that dramatically decreases from basic research through applied, coming to zero in the development phase. Dominant design industries are: food, textile, pulp and paper, printing and publishing, oil and paints, petroleum and coal, rubber, ceramics, iron and steel, transportation, energy. High tech industries are characterised by a freezing rate that decreases throughout the R&D process, but remains greater than zero even in the late development. High Tech industries are: drugs and medicines, ordinary machinery, electrical machinery, communications and electronics, precision equipment. Science based industries are those in which the freezing rate remains always high throughout the all R&D process. Science based industries are basically industrial chemicals. We adopted this taxonomy because it seems to be the most adequate for our future research about the role of contextual factors on the PMS design. Literature suggests in fact that the level and the type of uncertainty that characterises R&D activities significantly influences the PMS design [Shenhar and Dvir, 1996].

7 As far as company size is concerned, non-response bias was evaluated contrasting small and medium firms on one hand with large companies on the other.

8 The follow-up telephone interviews confirmed that in this latter case the “efficiency” is not used to measure the capacity of the single researcher or engineer to perform specific tasks; rather, it evaluates the resource consumption levels and respect of temporal milestones by R&D teams and projects.

9 In effect, by analyzing the relationships, a third model seems to emerge which is used in association with the “hard” and “motivational” archetypes. It could be called “soft” and it is aimed at improving coordination and communication, reducing uncertainty in R&D and favoring learning, through the use of efficiency and innovation capability performance dimensions (the same performance dimensions used by the other models). This is consistent with the theory advanced by Simons (2000) about the relationship between diagnostic and interactive control systems, with the latter requiring the use of the same performance dimensions, albeit in a different (interactive) fashion. Anyway, this hypothetical third model needs some further research, by means of a more extended survey and some in-depth case studies, to be better understood.

10 For each R&D performance measurement model we considered only the most relevant objectives and performance dimensions as suggested by the correlation coefficients reported in Table 3 in Appendix 3.

11 If company carries out both Research and Development activities, the filling of a questionnaire for each area of activity is required.

12 Uncertainty is intended here as the difference between the amount of information needed for successfully perform a specific activity and the information actually available. In this perspective, the measurement provides for systematic information and thus reduces uncertainty. Uncertainty could be financial, technical, commercial.

13 Learning is intended here as an improvement in the knowledge of the company’s R&D activities and of the external technological and market context. The measurement provides for systematic information and thus may stimulate learning

14 Numeric metrics obtained from the application of a definite algorithm that brings to the same evaluation independently from the person responsible for the measurement (e.g. percentage of projects concluded on time, number of citations of company’s researchers publications);

15 Metrics not expressed numerically, but through the personal judgment of the evaluator.

16 Internal reference standards are, for example, the historical internal data or the standard defined “ad hoc” for a specific performance/indicator (i.e. future objectives for R&D performance that reflect the overall business strategy).

17 External reference standards are industry/main competitors benchmarks.

18 Uncertainty is intended here as the difference between the amount of information needed for successfully perform a specific activity and the information actually available. In this perspective, the measurement provides for systematic information and thus reduces uncertainty. Uncertainty could be financial, technical, commercial.

19 Learning is intended here as an improvement in the knowledge of the company’s R&D activities and of the external technological and market context. The measurement provides for systematic information and thus may stimulate learning