J Manag Control (2011) 22:107–128
DOI 10.1007/s00187-011-0130-y

ORIGINAL PAPER

Using performance measures conceptually in innovation control

Sebastian Janssen · Klaus Moeller · Marten Schlaefke

Published online: 13 August 2011

Abstract In recent years, companies have been changing their innovation strategies, as they have realized that original new products can offer a major competitive advantage. Therefore, many companies are focusing on a more closely managed product-innovation process, and have consequently increased their use of performance-management frameworks. By doing so, such companies hope to increase the effectiveness and efficiency of their new-product-development activities. Within the performance-management framework, the use of innovation metrics plays an important and beneficial role.

For this reason, the present paper investigates the relationship between the design of innovation performance management frameworks and the actual utilization of information obtained from the implemented innovation metrics. We collected data from 133 technology-intensive companies and employed structural equation modeling for empirical analysis. This method allowed us to determine which design factors positively affect the extent to which managers conceptually use innovation metrics. In particular, we investigated how the balance, coherence, adaption, and user know-how of innovation metrics relate to their conceptual use.

Our results suggest that balance and user know-how of the metrics improve conceptual use of the performance measures, whereas no effect can be observed regarding coherence and adaption of the metrics. Thus, it seems highly advisable for firms

Dr. Sebastian Janssen is a former doctoral candidate of Prof. Dr. Klaus Moeller.
Prof. Dr. Klaus Moeller is Professor of Performance Management/Management Control and director of the Institute of Accounting, Controlling and Auditing at the University of St. Gallen, Switzerland and chairs the CEPRA—Center for Performance Research & Analytics, Augsburg.
Marten Schlaefke is research assistant at the chair of Performance Management/Management Control at the University of St. Gallen and a doctoral candidate of Prof. Dr. Klaus Moeller.

S. Janssen · K. Moeller (✉) · M. Schlaefke
Institute of Accounting, Control and Auditing, University of St. Gallen, Rosenbergstrasse 52, 9000 St. Gallen, Switzerland
e-mail: Klaus.Moeller@unisg.ch

to implement a simple, comprehensible performance management framework, consisting of financial and nonfinancial performance measures.

Keywords Performance measures · Innovation control · Performance management systems · Conceptual use

1 Introduction

Innovative products play an important role in building competitive advantage and can contribute significantly to companies' growth and profitability. Nevertheless, it is far from certain whether innovative ideas can transform into successfully launched products that meet their set financial targets (Cooper et al. 2004). A significant body of literature has therefore explored the relationship between the various potential success factors in product innovation and the performance outcomes (Balachandra and Friar 1997; Ernst 2002; Henard and Szymanski 2001; Montoya-Weiss and Calantone 1994). Currently, it is commonly accepted that successful product innovation depends heavily on such intangible elements as creativity and risk-taking behavior. However, the increased pace of innovation, shortened product life cycles, rapid advances in technology, and the globalization of new product development projects increasingly pressurize research and development to improve effectiveness and efficiency in new product development activities. Companies have therefore moved from a strategy of hope that—given the right mix of brains, money, equipment, and time to pursue ideas—scientists and engineers will eventually produce a profitable new product, to a more closely managed product innovation process strategy (Kerssens-van Drongelen et al. 2000; Wheelwright and Clark 1992).

Firms increasingly rely on formal management systems to manage new product development activities (Davila et al. 2004), because new product development performance can determine not only a firm's overall success and competitive advantage, but its very survival as well (Loch et al. 1996; Utterback 1994). Within the formal systems used to manage product innovation, innovation metrics play an important role (Davila 2000; Driva et al. 2000; Frattini et al. 2006; Kerssens-van Drongelen et al. 2000; Loch and Tapper 2002; Meyer et al. 1997). Yet it is widely recognized that innovation performance management is a complex task, since effort levels are difficult to observe, while success is highly uncertain, typically influenced by unmanageable factors, and can be assessed only after long delays (Frattini et al. 2006; Schumann and Ransley 1995). Furthermore, innovation performance management efforts can discourage high-level professionals' creativity and motivation as a result of the act of measuring (Brown and Svenson 1988). Firms are thus confronted with the question, how should innovation metrics be designed to support their beneficial use within the innovation process?

Consequently, the present study aims to broaden our understanding of the relationship between the design of innovation performance management frameworks and the actual utilization of the information obtained from the implemented innovation metrics. The term "information utilization" refers to the way managers employ innovation metrics for certain purposes within the product development context. Since the concept of information utilization is very broad, we will focus on the conceptual use of information, i.e., the application of information to develop and enlarge the managerial knowledge base. This method of using accounting information has been found to be beneficial with regard to general accounting information (Pelz 1978; Burchell et al. 1980), especially in the R&D context (Godener and Söderquist 2004; Cooper and Edgett 2008). We collected data from 133 technology-intensive companies and employed structural equation modeling to empirically analyze which design factors positively affect the extent to which managers use innovation metrics conceptually. In particular, we investigated how the balance, coherence, adaption, and user know-how of innovation metrics relate to their conceptual use.

Our findings suggest that some of the performance measurement design factors strongly influence the use of performance measures. The balance and user know-how of metrics improve the conceptual use of performance measures significantly, while no noteworthy effect can be observed regarding the coherence and adaption of metrics. Hence, innovation performance management frameworks need to be designed in accordance with the intended purpose of the measurement results. Our findings not only show the importance of the relationship between the design of innovation metrics and their conceptual use, but also underline the value of empirically validating the design factors of formal performance management systems within product innovation.

The remainder of the paper is organized as follows. In the next section, we examine the relevance of innovation control and innovation performance management frameworks. Furthermore, within the context of innovation, we discuss the information's usage form and exemplify a holistic framework for innovation control. Thereafter, we develop a theoretical model of the relationship between the design factors and conceptual use of metrics, and we present a set of hypotheses. After describing the survey design, we explain the data analysis by means of structural equation modeling. In the final section, we discuss theoretical contributions, practical implications, limitations, and insights for future research.

2 Innovation performance management

2.1 The relevance of innovation control

The term innovation signifies the "creative definition, development, and commercialization of substantially new products, services or businesses" (Davila et al. 2004, p. 28). A great many of articles underlines how important innovation is in providing firms with a competitive advantage. Recent decades, however, have seen a substantial change in the way innovations are managed and controlled within companies. Roussel et al. described the previous approach to innovation management as follows: "R&D is an overhead cost, a line item in the general manager's budget. General management participates little in defining programs or projects; funds are allocated to cost centers; cost control is at aggregate levels. There is minimum evaluation of the R&D results other than by those involved in R&D. The R&D activity is relatively isolated and there is little communication from R&D other than to say: Everything is going fine. There is only a modest sense of urgency: Things are ready when they are ready" (Roussel et al. 1991, p. 26). Innovation activities were considered creative but unstructured processes, and were "difficult, if not impossible, to manage and control" (Frattini et al. 2006, p. 426). Views on new product development have changed, although the importance of intangible elements, such as creativity and risk-taking behavior, is still widely acknowledged. Innovation activities are now understood as a largely repetitive process that can be managed and controlled. Hence, innovation management has moved from its previous hope-based strategy toward a more closely managed process (Davila et al. 2004).

The increased pace of innovation, shortened product life cycles, rapid advances in technology, and the ongoing globalization of new product development projects, all of which increasingly pressurize R&D to improve the effectiveness and efficiency of new product development activities, drive this development (Kerssens-van Drongelen et al. 2000; Frattini et al. 2006; Cardinal 2001). Consequently, firms need to manage and control their R&D function as intensely as they do their other functions by increasingly using formal management systems (Davila et al. 2004). Innovation metrics play an important role in such formal management systems, since they are a commonly accepted means of increasing the effectiveness and efficiency of organizational actions (Bremser and Barsky 2004; Cooper and Edgett 2008; Driva et al. 2000; Frattini et al. 2006; Kerssens-van Drongelen et al. 2000; Loch and Tapper 2002; Meyer et al. 1997).

2.2 Innovation performance management frameworks

Though according to Adams et al. "measurement of innovation management appears to be undertaken infrequently, in an ad hoc fashion, and relies on dated, unbalanced or under-specified models of the innovation management phenomenon" (Adams et al. 2006, p. 38), it is recommended that innovation metrics should be implemented as part of a systematic innovation management framework (Davila et al. 2005; Driva et al. 2000; Cooper and Edgett 2008; Kerssens-van Drongelen and Cook 1997; Möller et al. 2011a, 2011b). In recent years, companies have been developing more sophisticated performance management frameworks to provide relevant information to decision makers. Such frameworks can be used to capture and evaluate performance data and to identify key success factors within the innovation process. These frameworks can therefore be used to illustrate an organization's essential means and ends (Bourne et al. 2000; Garengo et al. 2005; Broadbent and Laughlin 2009).

Performance management frameworks can furthermore be used to link different innovation metrics to measure the efficiency and effectiveness of R&D activities. Such an innovation performance management framework can order innovation metrics hierarchically [e.g., the R&D return framework by Foster et al. (1985)], on the basis of a balanced scorecard approach (e.g., Bremser and Barsky 2004; Kerssens-van Drongelen and Cook 1997; Sandstrom and Toivanen 2002), or within a process-based, input-process, output-outcome framework (e.g., Brown and Svenson 1988; Möller and Janssen 2009). Moreover, the literature presents various

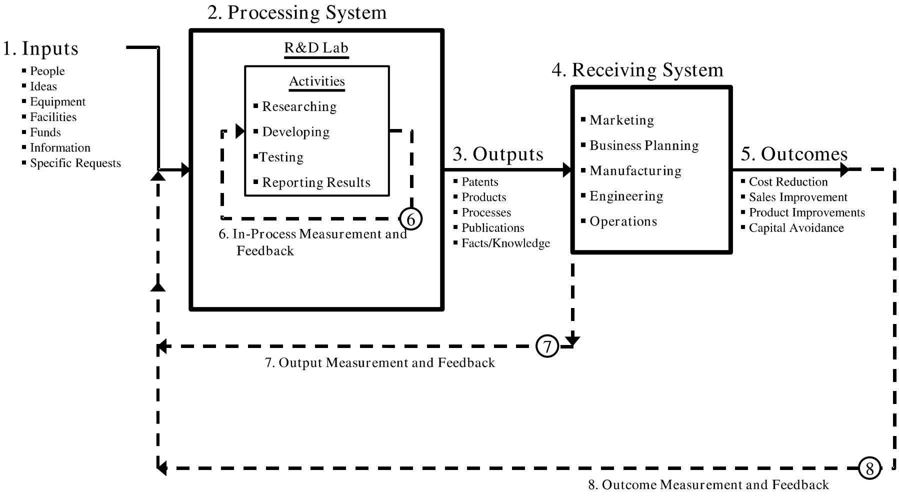


Fig. 1 Input-process-output-outcome framework (Brown and Svenson 1988, p. 12)

combinations of these frameworks (Chiesa et al. 2008; Codero 1990; Collins and Smith 1999; Davila et al. 2005; Godener and Söderquist 2004; Lee and Son 1996; Meyer et al. 1997; Muller et al. 2005; Ojanen and Vuola 2006; Pillai et al. 2002; Schumann and Ransley 1995). In the present study, we will focus on the input-process, output-outcome framework since, as noted by Kerssens-van Drongelen et al., it has been "widely used, though sometimes slightly adapted, in R&D management research" (Kerssens-van Drongelen et al. 2000, p. 115). This framework, depicted in Fig. 1, comprises the following four perspectives.

Input metrics measure the resources provided for R&D activities, for example, personnel, funds, equipment, and ideas. The largest cost drivers in new-product development are often the personnel and equipment costs. These factors can be measured by using a wide variety of absolute or relative, quantitative or qualitative metrics (Werner and Souder 1997; Loch et al. 1996). Intangible inputs, such as ideas, information, and know-how, are also of great importance to R&D activities (Hauser and Zettelmeyer 1997; Lee and Son 1996). However, some measurements of metrics, such as quality and experience of staff, are rather complicated compared with measuring the cost of immaterial inputs using metrics such as "expenditure on further education." Metrics that measure the financial side of R&D inputs are used widely in firms (Driva et al. 2000), although input metrics alone are of limited significance since an increase in input does not necessarily lead to an increase in output. Consequently, what inputs are used for (effectiveness) and how (efficiency) they are used within the innovation process are more important than their quantity. Relative metrics that combine input measures with process, output, and outcome measures thus provide valuable insight into R&D activities, especially when used for benchmarking. Moreover, an analysis of input metrics over time can help identify the trends and developments within R&D activities. Firms are therefore strongly advised to measure their inputs.

Process metrics measure the achievement of time, cost, and quality objectives as well as the project progress (Griffin and Page 1993; Driva et al. 2000; Kerssens-van Drongelen and Cook 1997). These metrics can be measured on the project level, on more aggregated levels, such as product lines or business units, or in respect to all R&D activities. The implementation of process measures helps identify deviations from plans at an early stage in the innovation process, allowing for quick corrective actions. Moreover, process metrics help develop more realistic assumptions and plans for future projects since they reflect lessons learned from other projects (Werner and Souder 1997). Project measures therefore keep experiences alive, improving the achievement of objectives based on past lessons learned.

Output metrics measure the direct results of R&D activities, i.e., new products or generated knowledge, and help identify trends and developments over time. These metrics can be absolute, for example, number of patents, or relative, for example, percentage of new products in the product program as a whole (Frattini et al. 2006). With regard to input measures, output metrics, such as "number of new products per dollar spent on R&D," can deliver valuable insights for improving the efficiency of the R&D process.

Outcome metrics measure innovation success in the market and thus focus on revenue, profit, market share, and customer satisfaction. The measurement of outcome metrics underlines innovations' economic importance for firms (Cooper and Edgett 2008). Absolute outcome metrics would be, for instance, revenue growth or market share growth due to new products. Here, the time in which a product is defined as new depends on companies' business models and can differ according to the industry. Relative outcome metrics are often measured in relation to inputs, such as sales per R&D employee or annual sales/R&D budget. The advantage of measuring outcome metrics is that it provides insight into the economic effects of R&D activities. However, depending on the time lag between development and market launch, it is mostly too late for corrective action once the outcome can be measured. Furthermore, an excessive focus on outcome measures can trigger myopia in the R&D department, thereby jeopardizing long-term projects (Hauser and Zettelmeyer 1997). Moreover, the financial success of innovative products not only depends on R&D activities, but is also strongly influenced by production, marketing, and sales. Outcome metrics nevertheless stress the importance of the economic effects of R&D activities and can help companies focus on economically successful projects.

Although the literature offers a wide variety of innovation metrics, innovation-performance management is considered a complex task: Effort levels are difficult to observe, while success is highly uncertain, is typically influenced by unmanageable factors, and can only be assessed after long delays (Frattini et al. 2006; Schumann and Ransley 1995). Furthermore, innovation performance management can discourage high-level professionals' creativity and motivation by the very act of measuring (Brown and Svenson 1988). Hence, the literature needs to focus on managers' use of metrics in order to develop empirically validated recommendations regarding how to successfully implement innovation performance management systems.

2.3 Usage forms of information in the innovation context

A literature review of the different ways of using information reveals great diversity in conceptualizing and classifying information utilization (Simon et al. 1954; Pelz 1978; Burchell et al. 1980; Feldman and March 1981; Ansari and Euske 1987; Menon and Varadarajan 1992; Hirst and Baxter 1993; Simons 1994; Vandenbosch 1999; Schäffer and Steiners 2004). In the present study, we follow the classification by Pelz (1978), who builds on work by Knorr (1977) and Rich (1977) to distinguish between an instrumental, symbolic, and conceptual way of using information. This classification has been successfully expanded to business administration and is often applied in this field (Diamantopoulos and Souchon 1999; Menon and Varadarajan 1992; Moorman 1995; Souchon and Diamantopoulos 1996; Souchon et al. 2003; Toften and Olsen 2004).

Instrumental use refers to applying information directly to solve a certain problem, while symbolic use denotes a political way of using information to legitimize decisions. Conceptual use indicates that information is applied to develop and enlarge the managerial knowledge base (Menon and Varadarajan 1992; Souchon et al. 2003). Hence, using innovation metrics conceptually relates to applying performance measures for learning purposes, which is a crucial R&D performance measurement goal (Kerssens-van Drongelen and Bilderbeek 1999; Godener and Söderquist 2004; Chiesa and Frattini 2007; Chiesa et al. 2009): "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 organizational improvement processes and strategic adaption" (Kerssens-van Drongelen and Bilderbeek 1999, p. 44). Godener and Söderquist (2004, p. 208) similarly argue that: "At the function level, performance metrics can be used for improving the measurement process itself, for improving the overall adherence to the research and new product development (R&NPD) process and for improving the R&NPD process as a whole." Cooper and Edgett (2008, p. 54) describe the importance of using innovation metrics for learning purposes even more dramatically: "Without metrics in place, project teams cannot be held accountable for results, while learning and continuous improvement are next to impossible."

We will, therefore, analyze the interplay between using innovation metrics conceptually and the performance management system design. But before we describe empirical testing of the relationships between design factors and conceptual use of metrics, we first present a performance management framework for R&D activities.

2.4 Providing a framework for innovation control

By taking the theoretical requirements for the conceptual use of innovation metrics into consideration, we have developed a performance management framework that combines these requirements comprehensively and is application oriented (see Fig. 2, Möller et al. 2011b). The framework can help to develop and design measurement systems that need to fulfill the criteria of a more intensive conceptual use of innovation metrics. Furthermore, using the framework will increase an awareness of such systems' important characteristics.

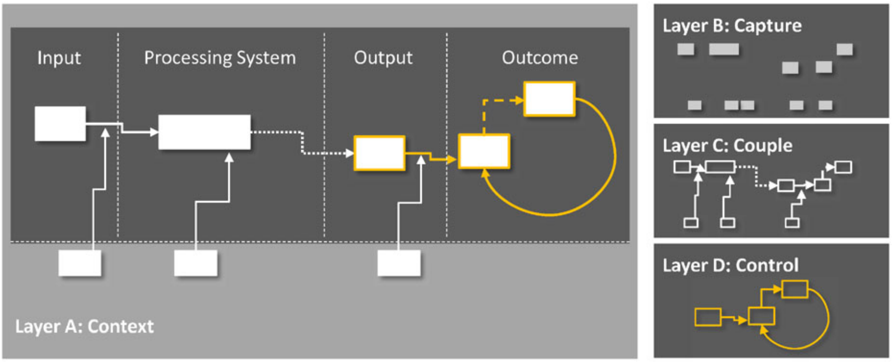


Fig. 2 Multilayer performance management framework

This framework could assist managers in evaluating and designing innovation metrics as well as in deciding whether these metrics have potential value for the new product development process. Furthermore, the framework could help managers obtain a better understanding of the innovation metrics that are used by clarifying the measurement system's crucial characteristics. Managers' know-how of the innovation metrics can therefore be increased by using this systematic framework to develop a measurement system.

The proposed framework consists of four different layers, which describe a performance measurement system's central features. The context layer considers internal and external factors that influence the organization and its processes. The second layer comprises the capturing of tangible and intangible performance drivers within the input, process, output, and outcome categories. The important requirements of innovation metrics should be considered at this point. By providing a complete overview of the existing metrics, the proposed framework forms the basis for a balanced and adapted metrics set. The third layer comprises the coupling of performance drivers. Coherent coupling can be done by proposing causal, final, and logical relationships between the different performance measures, by considering dependencies, time lags, etc. The fourth layer uses the knowledge generated by the previous layers to learn from previous actions, coordinate management actions, and regularly adjust the design of the performance measurement system. Since feedback and feedforward loops are used, the performance drivers' links have to be revised regularly by stimulating organization-wide continuous learning, and they can be used for what-if analysis and to test scenarios. By providing clear performance driver categories and, thus, a better overview of performance measures, the framework makes it possible for a well-balanced set of metrics to be designed. Managers can thus keep track of the new product development process. The framework addresses such characteristics as the balance, coherence, and adaption of performance measures, which can be the key for the conceptual use of innovation metrics. To date, there has been no empirical test of these characteristics' relevancy for innovation performance measurement frameworks. Consequently, we test these characteristics and provide evidence for the crucial characteristics.

3 Empirical identification of design factors that influence the conceptual use of performance measures

3.1 Theoretical framework

3.1.1 Research model

In this research, we focus on the interrelationship between accounting information and its actual use by innovation managers. Hence, we draw from behavioral accounting theory, which, as noted by Bruns and DeCoster (1969, p. 3), "considers the impact of the process of measuring and reporting on people and organizations," in two ways. First, we argue that managers use accounting information for different purposes. Of these different ways of use, we focus on the conceptual use of innovation metrics. Second, we argue that the extent to which innovation metrics are used conceptually depends on the design of the innovation performance management framework, and thus we focus on the balance and coherence of the performance management framework. Furthermore, we analyze the effect that performance measures and the periodic adaption of user know-how has on the conceptual use of performance measures. From an analysis of the literature, we argue that each of these factors could affect the extent to which innovation metrics could be used conceptually. Furthermore, since it seems likely that all the factors are interrelated, they should not be analyzed in isolation. We therefore employ structural equation modeling since factor analysis cannot analyze the interdependent factors sufficiently (Hair et al. 2010). In the following sections, we propose research hypotheses that will be tested later using structural equation modeling.

3.1.2 Hypotheses

Balance While traditional performance measurement systems focus only on the financial aspects of performance, modern multidimensional frameworks integrate nonfinancial performance measures to give an early indication of future business performance (Neely et al. 2007; Bourne et al. 2000). Nonfinancial measures play an important role, especially within new product development as financial measures are usually inappropriate for decision making owing to the time-lag problem. By the time financial information on the market success of product innovations becomes available, it is often too late for corrective action (Kerssens-van Drongelen et al. 2000). Hence, an innovation performance management system should consist of a balanced set of financial and nonfinancial performance measures to evaluate the success of innovation activities (future) (Driva et al. 2000; Godener and Söderquist 2004). A well balanced set of performance measures therefore provides a better view of the relevant aspects in the innovation process and enables the user to keep track of the progress of new product development activities. Furthermore, implementing a comprehensive set of performance measures helps managers divide their attention between different organizational objectives (Ullrich and Tuttle 2004), and it influences their cognition and motivation, which in turn influence their performance (Hall 2008). The following positive relationship is therefore proposed:

H1: The balance of innovation metrics is positively related to their conceptual use.

Coherence Coherence refers to the extent of the numerical or logical connections between metrics; it allows for testing and mapping of the causality-based links of performance-relevant factors within business activities (Eccles and Pyburn 1992; Neely et al. 2007). In contrast to traditional stand-alone metrics, a coherent set of performance measures should provide detailed insights into interrelations and cause-effect relationships, especially within new product development activities. Various studies therefore propagate the implementation of a coherent set of innovation metrics (Cooper and Edgett 2008; Davila et al. 2005; Driva et al. 2000; Kerssens-van Drongelen et al. 2000; Loch and Tapper 2002). Causality-based links are very important, especially within new product development activities, since once the product has been launched on the market, it is often too late for corrective action (Kerssens-van Drongelen et al. 2000). Consequently, a coherent set of leading financial performance measures should enhance the managerial knowledge base. Furthermore, empirical studies show that performance management systems that provide insight into the causal relationship of business activities positively affect organizational learning and, ultimately, the company's performance (Chenhall 2005; Ittner 2008). We thus argue that innovation metrics should be coherent in order to provide useful insight into the causal effects within new product development activities. Therefore, we propose the following relationship:

H2: The coherence of innovation metrics is positively related to their conceptual use.

Adaption To ensure that a set of metrics fulfills management's information demand, a performance measure review is necessary whenever the company's strategic or operative directions, or its competitive environment, change. Therefore, the complete set of performance measures should be reviewed and revised periodically (Bourne et al. 2000). This adaption process ensures that the performance management system always fits the business activities and prevents performance measures from reflecting former priorities (Burney and Widener 2007). Hence, the periodic review of performance measures influences the usability of the information generated and is especially important in rapidly changing environments (Henri 2010; Malina and Selto 2004), such as new product development. Furthermore, Henri explicitly states that performance measures may be more effective in supporting education and learning throughout the organization if they are periodically revised (Henri 2010, p. 77). We therefore argue that adaption of metrics is an important determinant of their conceptual use and hypothesize:

H3: The adaption of innovation metrics is positively related to their conceptual use.

User know-how This factor measures the extent to which the user understands the calculation and meaning of the metrics. Though a company can control this factor only indirectly, its importance in the context of the use of innovation metrics is obvious. It seems unlikely that managers with only a little understanding of the meaning of innovation metrics could use them to enlarge their knowledge base; this phenomenon is known as the measurement-use gap (Stivers et al. 1998). In the present study, we will therefore analyze the effect of user know-how on the extent to which innovation metrics are used conceptually, and we propose the following positive relationship:

H4: User know-how of innovation metrics is positively related to their conceptual use.

3.1.3 Constructs

The *conceptual use* of innovation metrics describes applying performance measures for learning and continuous improvement, thus covering an important R&D performance measurement goal (Kerssens-van Drongelen and Bilderbeek 1999; Godener and Söderquist 2004; Chiesa and Frattini 2007; Chiesa et al. 2009). Godener and Söderquist (2004, p. 208) describe a typical situation in which innovation metrics were used conceptually. A company ascertained that its patents generated very little revenue in their first year and decided to investigate their patent writing process. A metric was used to assess the patent revenues, which led to the conclusion that the patents were often so poorly worded that the company's competitors could easily find loopholes to circumvent them. Subsequently, the company's entire patent-writing process was modified. Hence, the conceptual use of innovation metrics does not cause immediate decision making, but it serves to develop and enlarge the managerial knowledge base (Menon and Varadarajan 1992; Souchon et al. 2003). The measurement instrument builds on work by Moorman (1995) has been tested and validated in several empirical studies (Schäffer 2007). The Cronbach α of 0.86 obtained in this study indicates the items' high internal consistency in the measurement instrument.

The construct *balance* reflects measures managers' assessment of the balance between the financial and nonfinancial performance measures used to manage and control new product development activities. This construct yielded a Cronbach α of 0.84, thus indicating the scale's high internal reliability. *Coherence* measures managers' assessment of coherence in the metrics used to manage and control new product development activities, and it yielded a Cronbach α of 0.80. *Adaption* measures managers' assessment of the regular revision of the metrics used to manage and control new product development activities. The Cronbach α of 0.87 obtained in this study indicates the high internal consistency of the items in the measurement instrument (Sandt 2004). *User know-how* measures the extent to which managers perceive themselves as capable of understanding and utilizing the metrics to manage and control new product development activities. The achieved Cronbach α of 0.90 also indicates a high internal reliability (Sandt 2004; Frank 2000; Schäffer and Steiners 2004).

The wording of the construct items is provided in Table 1. To reduce response error, the questionnaire was built on tested and validated measures from prior studies (Schäffer 2007; Dillman 2007; Van der Stede et al. 2005). All measures are based on seven-point Likert-type scales.

3.2 Method and results

3.2.1 Data collection and sample

Data were gathered by means of a survey research method, which consisted of administering a written questionnaire to the innovation managers of German engineering companies (for details of the study see also Janssen 2011). The survey was carried out in the German engineering industry, because it is characterized by a high level of innovativeness (Aschhoff et al. 2008). Moreover, it was felt that the German engineering industry's strong globalization would enhance the results' generalizability. To ensure a minimum level of innovation management within the contacted firms, we restricted the sample to companies with more than 750 employees. A professional database was used, and 840 firms fulfilled the screening criterion.

We used a survey instrument to collect the data for the study. Questionnaire instruments documented in the extant academic literature as well as theoretical input from management control and innovation research were used as the basis for an initial draft. We undertook extensive pretests with teams of academic and industrial experts, using a retrospective and think-aloud technique (Dillman 2007). Based on the responses, we reordered some parts of the questionnaire and a number of items were subsequently reworded. Using a key informant approach (Kumar et al. 1993), we usually obtained data from the head of R&D. By limiting the study to senior innovation management, we assured that the respondents had a broad view of their companies' new product development activities and could provide the detailed management information required. We included several aspects of survey design to encourage a higher response rate, including a friendly, personalized covering letter explaining the purpose of the research study, consideration of the length of the survey, and the order in which the questions appeared (Bosnjak et al. 2005; Dillman 2007; Hague 1993).

Before we distributed the questionnaires, we checked the accuracy of the data on the 840 companies we obtained from the database. A total of 172 firms were excluded from the survey as they were either not engaged in new product development activities (they were only holding companies) or they could not be contacted (they had gone out of business). We contacted the remaining 668 companies by phone between May and August 2009 in an attempt to identify the R&D heads. In total, 198 firms refused to reveal the identity of the head of R&D, mostly citing compliance with internal company guidelines. The remaining 470 identified heads of R&D were then contacted by phone and asked to participate in the survey. The questionnaire was e-mailed to the interested heads and could be returned by regular mail, e-mail, or fax. To enhance the response rate, we made follow-up telephone calls and promised the participants an executive summary of the findings as well as an individual benchmark report for their company. A total of 133 sufficiently completed questionnaires was received, thus yielding a response rate of 28.3 percent.

Nonresponders often differed significantly from responders (De Vaus 2002). We therefore tested for nonresponse bias, based on a comparison of early and late responders (Armstrong and Overton 1977). Structural data, such as the companies' number of employees, sales volume, and R&D investment, showed no significant difference between the two groups. Furthermore, early and late responders showed no significant difference regarding any of the construct items. Thus, the results confirm the data's representativeness in respect of German engineering companies that have more than 750 employees. The sample firms had an average of 10,657 employees (median 2,200) and revenue of 2,300 million EUR (median 380 million). On average, the companies invested 4.5 percent of their revenues in R&D (median 3.5 percent) and successfully completed 21 new product development projects per year (median 6). Almost 90 percent of the respondents were the head of R&D or a senior innovation manager with an average tenure in the position of 8.5 years (median 7 years). Hence, the firms, as well as respondents in the sample, were highly suitable for this research.

On average, 2.8 percent missing values were observed in the construct-measurement scale items. Missing values were imputed by means of the expectation-maximization (EM) algorithm prior to the main data analyses (Bernaards and Sijtsma 1999).

As it seemed likely that all the factors were interrelated, we employed structural equation modeling (SEM) (Bollen 1989; Kline 2005) to estimate the multivariate dependencies that the research hypotheses addressed. SEM combines the multivariate technique factor analysis and multiple regression analysis, thus allowing the estimation of multiple and interrelated dependence relationships as well as the representation of unobserved theoretical constructs (latent variables), while directly accounting for measurement error in the estimation (Hair et al. 2010).

In accordance with Kline (2005), we applied a two-step SEM procedure. In the first step, a confirmatory factor analysis (CFA) was undertaken to assess the construct measurement. In the second step, we specified and evaluated a structural model, as implied by the research hypotheses.

3.2.2 Measurement model

We examined the normality of each variable before evaluating the construct measurement and overall model fit. According to West et al. (1995), the absolute value of a skewness index of less than 2 and a kurtosis index of less than 7 indicate normal distribution. The analysis results showed that all of the items' skewness and kurtosis indices were less than the recommended threshold values. However, the test results of Mardia's coefficient (Mardia 1970) indicated multivariate kurtosis. Consequently, we decided to use the maximum likelihood estimation (ML) procedure using AMOS 17.0 software (Arbuckle 2006) to estimate the model parameters and test the proposed model's adequacy. The parameter estimates generated by the ML were found to be relatively robust against non-normality (Hair et al. 2010; Kline 2005). However, the results of statistical tests were potentially positively biased and the standard errors tended to be too low (West et al. 1995). Thus, following Byrne (2001), we also assessed the parameter estimates' significance by means of bias-corrected bootstrapping (Efron and Tibshirani 1993).

The measurement model's fit was assessed by means of local and global fit statistics. Local fit statistics were used to evaluate the construct reliability and validity (Bagozzi and Baumgartner 1994; Hair et al. 2010) and whether the constructs were sufficiently distinguishable (Fornell and Larcker 1981). Global fit statistics were employed to assess the overall model fit, i.e., whether the model appropriately reproduced the empirical associations between the variables (Boomsma 2000; Kline 2005).

The global fit statistics, reported in Table 1, indicate a good model fit. A marginal violation was observed in respect to the adjusted goodness-of-fit index (AGFI). The AGFI, however, is known to be of only limited statistical power (Sharma et al. 2005).

Table 1 Measurement scales and measures of local fit for the CFA model

Construct and measurement item	Factor Reliability	Average Variance extracted	Std. Factor Loadings	ML p-value
Threshold for acceptable fit[a]	= 0.6	= 0.5	= 0.5	= 0.05
Conceptual use[b]	0.86	0.67		
Through the metrics I gain insights into our NPD activities' interrelationships.			0.761	<0.001
The metrics enlarge my knowledge base of our NPD activities' general situation.			0.903[c]	
The metrics enhance the assessment of our NPD activities' trends and future developments.			0.806	<0.001
Balance of metrics[b]	0.86	0.67		
The innovation metrics comprise non-financial metrics (e.g., number of patents, success rate of innovation projects, customer satisfaction, etc.) in addition to financial metrics.			0.782[c]	
The innovation metrics take the quantitative and qualitative perspectives of new product development activities into account.			0.871	<0.001
The innovation metrics cover both the "hard" and "soft" aspects of new product development activities.			0.803	<0.001
Coherence of metrics[b]	0.80	0.67		
The innovation metrics are highly coherent.			0.881[c]	
The innovation metrics amount to an innovation performance measurement system			0.773	<0.001
Adaption of metrics[b]	0.87	0.70		
The metrics are revised periodically to fit to new requirements.			0.812[c]	
When our new product development activities are changed substantially, we review the expediency of our innovation metrics.			0.910	<0.001
The innovation metrics are revised when organizational changes occur (e.g., new strategy, new technology, etc.)			0.786	<0.001
User know-how[b]	0.90	0.75		
I understand the calculation of the metrics fully.			0.926[c]	
I don't need to reflect on the metrics' meaning.			0.808	<0.001
I could easily explain the metrics to a third party.			0.855	<0.001

Notes: [a]For the thresholds of acceptable local fit statistics see Bagozzi and Yi (1988), Hair et al. (2010), Kline (2005). Global fit statistics for the CFA model with threshold values in parentheses (Baumgartner and Homburg 1996; Hair et al. 2010; Kline 2005): $\chi^2/df = 1.59$ (<3), RMSEA = 0.07 (= 0.05–0.10), GFI = 0.90 (= 0.90), AGFI = 0.85 (= 0.90), CFI = 0.97 (= 0.90), TLI = 0.95 (= 0.90), SRMR = 0.04 (= 0.05). [b]Seven-point Likert scale: 1 strongly disagree, 7 strongly agree, [c]fixed parameter

Table 2 Fornell-Larcker criterion

		1	2	3	4	5
1	Conceptual use	(0.67)				
2	Balance of metrics	0.41	(0.67)			
3	Coherence of metrics	0.43	0.45	(0.67)		
4	Adaption of metrics	0.31	0.29	0.52	(0.70)	
5	User know-how	0.27	0.08	0.32	0.09	(0.75)

Moreover, the root mean square error of approximation (RMSEA) indicates an acceptable model fit, while the chi-square statistic, adjusted for the degrees of freedom (χ^2/df), the goodness-of-fit index (GFI), the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the standardized root mean residual (SRMR) indicate a good overall model fit (Baumgartner and Homburg 1996; Hair et al. 2010; Kline 2005).

The results of the local fit measures indicated that the construct measurement was highly reliable and valid. We assessed the constructs' reliability by means of indicator and factor reliability (Bagozzi and Yi 1988). Though the factor reliability can also be used to assess the constructs' convergent validity, it is recommended that the average variance extracted (AVE) and the factor loadings' size and significance be included when assessing the convergent validity (Hair et al. 2010).

We furthermore conducted the Fornell-Larcker test to ascertain the constructs' discriminant validity, which requires each construct's average variance extracted (AVE) to be larger than the squared correlation between all pairs of factors in the model (Fornell and Larcker 1981). The results, displayed in Table 2, indicated adequate discriminant validity.

3.2.3 Structural model

Having ensured high measurement quality, we specified the structural relationships between the constructs, as implied by hypotheses H1 to H4, in a second step. The structural model yielded almost the same global fit statistics as the CFA model (see Table 3), indicating good overall model fit.

As shown in Table 3, the model strongly supports hypotheses H1 and H4, whereas hypotheses H2 and H3 are not supported. Specifically, no evidence was found that either the metrics' coherence or their adaption had a significant impact on their conceptual use. Thus, the model does not support hypotheses H1 and H3. Hypothesis H1 posits a direct positive relationship between the balance of innovation metrics and their conceptual use. As shown in Fig. 2, there is a significant positive relationship ($p \leq 0.001$) between the balance of metrics and their conceptual use. This indicates support for H1. H4 states that user know-how increases the extent to which innovation metrics are used conceptually. The results show a significant positive relationship ($p \leq 0.01$) between user know-how and the conceptual use of metrics, thus supporting H4. We also assessed the significance of parameter estimates by means of bias-corrected bootstrapping. This procedure supported our results.

Table 3 Results of the path analysis

Structural path	Maximum likelihood	Bootstrap
H1: Balance & Conceptual use	0.41***	0.41**
H2: Coherence & Conceptual use	0.05^{ns}	0.05^{ns}
H3: Adaption & Conceptual use	0.21^{ns}	0.21^{ns}
H4: User know how & Conceptual use	0.32**	0.32*

Note: Global fit statistics with threshold values in parentheses (Baumgartner and Homburg 1996; Hair et al. 2010; Kline 2005): $\chi^2/df = 1.59$ (<3), RMSEA = 0.07 (= 0.05–0.10), GFI = 0.90 (= 0.90), AGFI = 0.85 (= 0.90), CFI = 0.97 (= 0.90), TLI = 0.95 (= 0.90), SRMR = 0.04 (= 0.05)

$^{***}p = 0.001$, $^{**}p = 0.01$, $^{*}p = 0.05$, ^{ns}not significant

Table 4 The moderating effect of user know-how

Moderated path in the chi-square difference test	The chi-square difference test between low and high user know-how for the moderated path	
(balance, coherence, adaption) & conceptual	$\chi^2 = 5.894$; $df = 3$	Not significant ($p = 0.117$)
balance & conceptual	$\chi^2 = 1.665$; $df = 1$	Not significant ($p = 0.197$)
coherence & conceptual	$\chi^2 = 0.128$; $df = 1$	Not significant ($p = 0.720$)
adaption & conceptual	$\chi^2 = 2.153$; $df = 1$	Not significant ($p = 0.142$)

Furthermore, we tested for moderating effects (see Table 4), as it can be argued that user know-how moderates the effect that the balance, coherence, and adaption of metrics has on their conceptual use. To evaluate the moderating effects, user know-how was divided into high and low groups on the basis of a median split (Dabholkar and Bagozzi 2002). Subsequently, the core model was tested for high- and low-know-how groups by using structural equation modeling. We performed rigorous pretests of the measurement model invariance (Jöreskog 1971) to verify that the items comprising a particular measuring instrument functioned similarly across the groups (Kline 2005; Byrne 2001). In testing for measurement model invariance across the groups, sets of parameters were put to the test in a logically ordered and increasingly restrictive fashion (Byrne 2001). Once we had assured measurement-model invariance, we tested the moderating effect of user know-how, assuming equal factor loadings. As shown in Table 4, the χ^2 difference tests indicated that the observed differences between the path loadings were insignificant. Consequently, user know-how does not moderate the effect of balance, coherence, and adaption of metrics on their conceptual use.

4 Discussion

The results of our research underline the importance of a balanced set of innovation metrics since a balanced framework increases the extent to which innovation performance measures are used conceptually. Our research therefore empirically validates the call to increase the balance of innovation performance management frameworks by implementing quantitative and qualitative measures (Godener and Söderquist 2004). Moreover, our results are in line with the increasing implementation of multidimensional performance management frameworks that cover financial as well as nonfinancial aspects.

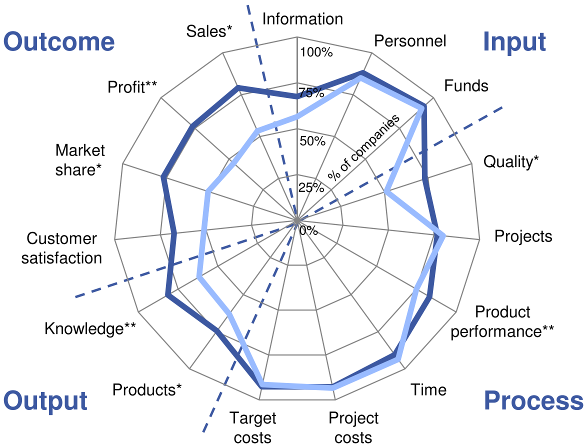
To our surprise, neither the coherence nor the adaption of innovation metrics has a significant impact on their conceptual use. Since the coherence of metrics and their regular adaption are widely acknowledged design factors, we believe that our results reflect the specifics of using innovation metrics. Most of the sampled companies had recently implemented innovation metrics to track R&D activities. Hence, the need to adapt these metrics might not yet have occurred, since most of these companies' innovation performance management systems were still up to date. Furthermore, using metrics conceptually, i.e., enlarging the knowledge base, might not necessarily require a coherent set of metrics, but could perhaps also be based on stand-alone metrics.

User know-how strongly influences the extent to which innovation metrics are used conceptually. This result seems to be fairly straightforward, and underlines the importance of implementing easily comprehensible performance measures. Furthermore, companies can trigger the conceptual use of innovation metrics by offering managers workshops and training to ensure that they understand the metrics. In addition, our results support the idea of lean performance management systems that do not necessarily need to cover a wide range of performance measures.

The results of the structural equation model underline the substantial impact that a balanced set of performance measures has on the conceptual use of innovation metrics. Nevertheless, from these results, we cannot ascertain how firms can improve the balance of their innovation performance measurement system. Hence, we analyzed the balance of innovation metrics in more detail on the basis of the four perspectives of an input-process-output-outcome framework. We aimed at developing practical guidelines to increase the balance within innovation performance measurement systems, focusing on the differences between successful and unsuccessful companies' implemented metrics.

To differentiate between successful and unsuccessful companies, we split the dataset into three equally-sized clusters. The most successful third was labeled "top performers," while the least successful third was labeled "low performers." The classification into top and low performers was based on two criteria, which the respondents assessed on a Likert-type scale ranging from 1 (very bad) to 7 (very good). The first criterion measures the average achievement of innovation projects' time, cost and quality objectives by means of six items. The second criterion measures product innovations' financial success and comprises three items. The respondents were asked to assess the average financial success of their company's new product development activities over the last five years by comparing their profitability, profit contribution, and overall profit effect with those of their main competitor. The first and second measure were then weighted equally and used as a compound measure of financial success to differentiate between top and low performers. The comparison of the mean of the top and low performers produced no significant differences between their revenue, number of employees, and R&D budget. Hence, the two groups are structurally comparable.

Fig. 3 Balance of innovation metrics



The balance of innovation metrics in the two groups was assessed by means of an input-process, output-outcome framework, and the results are displayed in Fig. 3 (Janssen and Möller 2011). The results show that more top performers than low performers measure immaterial inputs and process aspects focusing on quality and project progress. Moreover, significantly more top performers than low performers implemented output and outcome measures. These results are in line with those of previous studies, indicating that most companies can increase the balance in their set of innovation metrics by implementing more (financial) outcome measures (Kerssens-van Drongelen and Bilderbeek 1999; Donnelly 2000). Moreover, the results can be interpreted as being indicative of the beneficial impact of using innovation metrics conceptually, since top-performing firms foster the conceptual use of innovation metrics by implementing a balanced set of performance measures. Based on this analysis, we therefore recommend a stronger integration of (financial) outcome measures to increase the balance within the innovation measurement system and, thus, promote the conceptual use of innovation metrics.

Our research is subject to several limitations associated with the survey format and the use of structural equation modeling. First, we limited our research to the German engineering industry and companies with more than 750 employees. Future research should therefore cross-validate our findings with data from other industries and use different company sizes. Furthermore, our data are cross-sectional and employ a key informant approach, which is widely used (Page and Schirr 2008), but has inherent limitations. The contacted innovation managers can, however, be considered well-informed respondents with respect to innovation performance measurement and the performance of new product development activities. The study used perceptual performance measures as it is difficult to obtain objective data for innovation-related performance from German industry. Nevertheless, several studies show that there is a high correlation between perceptual and objective measures. Another limitation of our research is the aggregated view of innovation metrics, since we do not differentiate between the process phases in which these metrics are used and those innovation projects subject to measurement. Both aspects could, however, have an impact on the interrelationship between design factors and the conceptual use of metrics. Hence, future research could fruitfully analyze the interrelationship between design factors and the conceptual use of metrics on a less aggregated level.

To summarize, our results suggest that performance measurement design factors strongly influence the use of innovation performance measurement. The balance of metrics and user know-how improves the conceptual use of performance measures significantly, while no significant effect can be observed regarding the coherence and adaption of metrics. Thus, it seems highly advisable for firms to implement a simple, comprehensible performance measurement framework, consisting of financial and nonfinancial performance measures. In addition, our results underline the importance of empirically analyzing the interrelationship between the design and use of management accounting tools to develop validated performance measurement design recommendations for companies.

References

Adams, R., Bessant, J., & Phelps, R. (2006). Innovation management measurement: a review. *International Journal of Management Reviews*, *8*(1), 21–47.

Ansari, S., & Euske, K. J. (1987). Rational, rationalizing, and reifying uses of accounting data in organizations. *Accounting, Organizations and Society*, *12*(6), 549–570.

Arbuckle, J. L. (2006). *AMOS 7.0 user's guide*. Chicago.

Armstrong, J. S., & Overton, T. S. (1977). Estimating nonresponse bias in mail surveys. *Journal of Marketing Research*, *14*, 396–402.

Aschhoff, B., Doherr, T., Köhler, C., Peters, B., Rammer, C., Schubert, T., & Schwiebacher, F. (2008). *Innovation in Germany*. Centre for European Economic Research. ftp://ftp.zew.de/pub/zew-docs/mip/07/MIP_2007_engl.pdf.

Bagozzi, R., & Baumgartner, H. (1994). The evaluation of structural equation models and hypothesis testing. In R. Bagozzi (Ed.), *Principles of marketing research* (pp. 386–422), Cambridge, MA.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation on structural equation models. *Journal of the Academy of Marketing Science*, *16*, 74–94.

Balachandra, R., & Friar, J. H. (1997). Factors for success in R&D projects and new product innovation: a contextual framework. *IEEE Transactions on Engineering Management*, *44*(3), 276–288.

Baumgartner, H., & Homburg, C. (1996). Applications of structural equation modeling in marketing and consumer research: a review. *International Journal of Research in Marketing*, *13*, 139–161.

Bernaards, C. A., & Sijtsma, K. (1999). Factor analysis of multidimensional polytomous item response data suffering from ignorable item nonresponse. *Multivariate Behavioral Research*, *34*, 277–313.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York.

Boomsma, A. (2000). Reporting analyses of covariance structures. *Structural Equation Modeling*, *7*, 461–483.

Bosnjak, M., Tuten, T. L., & Wittmann, W. W. (2005). Unit (non)response in web-based access panel surveys: an extended planned-behavior approach. *Psychology & Marketing*, *22*(6), 489–505.

Bourne, M., Mills, J., Wilcox, M., Neely, A., & Platts, K. (2000). Designing, implementing and updating performance measurement systems. *International Journal of Operations & Production Management*, *20*(7), 754–771.

Bremser, W. G., & Barsky, N. P. (2004). Utilizing the balanced scorecard for R&D performance measurement. *R&D Management*, *34*(3), 229–238.

Broadbent, J., & Laughlin, R. (2009). Performance management systems: a conceptual model. *Management Accounting Research*, *20*(4), 283–295.

Brown, M. G., & Svenson, R. A. (1988). Measuring R&D productivity. *Research Technology Management*, *31*(4), 11–15.

Bruns, W. J., & DeCoster, D. T. (1969). *Accounting and its behavioral implications*. New York.

Burchell, S., Clubb, C., Hopwood, A., Hughes, J., & Nahapiet, J. (1980). The role of accounting in organization and society. *Accounting, Organizations & Society*, *5*(1), 5–27.

Burney, L., & Widener, S. K. (2007). Strategic performance measurement systems, job-relevant information, and managerial behavioral responses—role stress and performance. *Behavioral Research in Accounting*, *19*, 43–69.

Byrne, B. M. (2001). *Structural equation modeling with AMOS, basic concepts, applications, and programming*. New Jersey.

Cardinal, L. B. (2001). Technological innovation in the pharmaceutical industry: the use of organizational control in managing research and development. *Organization Science*, *12*(1), 19–36.

Chenhall, R. H. (2005). Integrative strategic performance measurement systems, strategic alignment of manufacturing, learning and strategic outcomes: an exploratory study. *Accounting, Organizations & Society*, *30*(5), 395–422.

Chiesa, V., & Frattini, F. (2007). Exploring the differences in performance measurement between research and development: evidence from a multiple case study. *R&D Management*, *37*(4), 283–301.

Chiesa, V., Frattini, F., Lazzarotti, V., & Manzini, R. (2008). Designing a performance measurement system for the research activities: a reference framework and an empirical study. *Journal of Engineering & Technology Management*, *25*(3), 213–226.

Chiesa, V., Frattini, F., Lazzarotti, V., & Manzini, R. (2009). Performance measurement in R&D: exploring the interplay between measurement objectives, dimensions of performance and contextual factors. *R&D Management*, *39*(5), 487–519.

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

Collins, J., & Smith, D. (1999). Innovation metrics: a framework to accelerate growth. In *Prism* (pp. 33–47).

Cooper, R. G., & Edgett, S. J. (2008). Maximizing productivity in product innovation. *Research Technology Management*, *51*(2), 47–58.

Cooper, R. G., Edgett, S. J., & Kleinschmidt, E. J. (2004). Benchmarking best NPD practices—I. *Research Technology Management*, *47*(1), 31–43.

Dabholkar, P. A., & Bagozzi, R. P. (2002). An attitudinal model of technology—based self—service: moderating effects of consumer traits and situational factors. *Journal of the Academy of Marketing Science*, *30*, 184–201.

Davila, T. (2000). An empirical study on the drivers of management control systems' design in new product development. *Accounting, Organizations & Society*, *25*(4/5), 383–409.

Davila, T., Epstein, M. J., & Matusik, F. (2004). Innovation strategy and the use of performance measures. *Advances in Management Accounting*, *13*, 27–58.

Davila, T., Epstein, M. J., & Shelton, R. (2005). *Making innovation work: how to manage it, measure it, and profit from it*. Philadelphia, PA.

De Vaus, D. (2002). *Surveys in social research*, 5th ed. London.

Diamantopoulos, A., & Souchon, A. L. (1999). Measuring export information use: scale development and validation. *Journal of Business Research*, *46*(1), 1–14.

Dillman, D. A. (2007). *Mail and Internet surveys: the Tailored design method*, 2nd ed. Hoboken, NJ.

Donnelly, G. (2000). A P&L for R&D. *CFO*, *16*(2), 44–50.

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

Eccles, R. G., & Pyburn, P. J. (1992). Creating a comprehensive system to measure performance. *Management Accounting*, *74*(4), 41–44.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York.

Ernst, H. (2002). Success factors of new product development: a review of the empirical literature. *International Journal of Management Reviews*, *4*(1), 1.

Feldman, M. S., & March, J. G. (1981). Information in organizations as signal and symbol. *Administrative Science Quarterly*, *26*(2), 171–186.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50.

Foster, R. N., Linden, L. H., Whiteley, R. L., & Kantrow, A. M. (1985). Improving the return on R&D—II. *Research Management*, *28*(2), 13–22.

Frank, S. (2000). *Erfolgreiche Gestaltung und Kostenrechnung: Determinanten und Wirkungen am Beispiel mittelständischer Unternehmen*. Wiesbaden.

Frattini, F., Lazzarotti, V., & Manzini, R. (2006). Towards a system of performance measures for research activities: Nikem research case study. *International Journal of Innovation Management*, *10*(4), 425–454.

Garengo, P., Biazzo, S., & Bititci, U. S. (2005). Performance measurement systems in SMEs: a review for a research agenda. *International Journal of Management Reviews*, *7*(1), 25–47.

Godener, A., & Söderquist, K. E. (2004). Use and impact of performance measurement results in R&D and NPD: an exploratory study. *R&D Management*, *34*(2), 191–219.
Griffin, A., & Page, A. L. (1993). An interim report on measuring product development success and failure. *Journal of Product Innovation Management*, *10*(4), 291–308.
Hague, P. N. (1993). *Questionnaire design*. London.
Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: a global perspective*, 7th ed. Upper Saddle River, NJ.
Hall, M. (2008). The effect of comprehensive performance measurement systems on role clarity, psychological empowerment and managerial performance. *Accounting, Organizations & Society*, *33*(2/3), 141–163.
Hauser, J. R., & Zettelmeyer, F. (1997). Metrics to evaluate R&D. *Research Technology Management*, *40*(4), 32–38.
Henard, D. H., & Szymanski, D. M. (2001). Why some new products are more successful than others. *Journal of Marketing Research*, *38*(3), 362–375.
Henri, J.-F. (2010). The periodic review of performance indicators: an empirical investigation of the dynamism of performance measurement systems. *European Accounting Review*, *19*(1), 73–96.
Hirst, M. K., & Baxter, J. A. (1993). A capital budgeting case study: an analysis of a choice process and roles of information. *Behavioral Research in Accounting*, *5*, 187.
Ittner, C. D. (2008). Does measuring intangibles for management purposes improve performance? A review of the evidence. *Accounting & Business Research*, *38*(3), 261–272.
Janssen, S. (2011). *Innovationssteuerung mit Kennzahlen – Erfolgswirkungen und Determinanten einer konzeptionellen Kennzahlennutzung*. Göttingen.
Janssen, S., & Möller, K. (2011). Erfolgreiche Steuerung von Innovationsprozessen und – projekten – Ergebnisse einer empirischen Studie. *Zeitschrift für Controlling und Management*, *55*(2), 96–103.
Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, *36*, 409–426.
Kerssens-van Drongelen, I. C., & Bilderbeek, J. (1999). R&D performance measurement: more than choosing a set of metrics. *R&D Management*, *29*(1), 35.
Kerssens-van Drongelen, I. C., & Cook, A. (1997). Design principles for the development of measurement systems for research and development. *R&D Management*, *27*(4), 345–357.
Kerssens-van Drongelen, I. C., Pearson, A., & Nixon, B. (2000). Performance measurement in industrial R&D. *International Journal of Management Reviews*, *2*(2), 111–143.
Kline, R. B. (2005). *Principles and practice of structural equation modeling*, 2nd ed. New York.
Knorr, K. D. (1977). Policymakers' use of social science knowledge: symbolic or instrumental. In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 165–182), Lexington.
Kumar, N., Stern, L. W., & Anderson, J. C. (1993). Conducting interorganizational research using key informants. *Academy of Management Journal*, *36*, 1633–1651.
Lee, M., & Son, B. (1996). Measuring R&D effectiveness in Korean companies. *Research Technology Management*, *39*(6), 28–31.
Loch, C. H., & Tapper, U. A. 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., & Terwiesch, C. (1996). Measuring development performance in the electronics industry. *Journal of Product Innovation Management*, *13*(1), 3–20.
Malina, M. A., & Selto, F. H. (2004). Choice and change of measures in performance measurement models. *Management Accounting Research*, *15*(4), 441–469.
Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*, 519–530.
Menon, A., & Varadarajan, P. R. (1992). A model of marketing knowledge use within firms. *Journal of Marketing*, *56*(4), 53–71.
Meyer, M. H., Tertzakian, P., & Utterback, J. M. (1997). Metrics for managing research and development in the context of the product family. *Management Science*, *43*(1), 88–111.
Möller, K., & Janssen, S. (2009). Performance Measurement von Produktinnovationen: Konzepte, Instrumente und Kennzahlen des Innovationscontrollings. *Controlling*, *21*(2), 89–96.
Möller, K., Menninger, J., & Robers, D. (2011a). *Innovationscontrolling*. Stuttgart.
Möller, K., Schläfke, M., & Schönefeld, C. (2011b). Wirkungsorientiertes Performance Management. *Controlling*, *23*(7), 372–378.
Montoya-Weiss, M. M., & Calantone, R. (1994). Determinants of new product performance: a review and meta-analysis. *Journal of Product Innovation Management*, *11*(5), 397–417.
Moorman, C. (1995). Organizational market information processes: cultural antecedents and new product outcomes. *Journal of Marketing Research*, *32*(3), 318–335.

Muller, A., Valikangas, L., & Merlyn, P. (2005). Metrics for innovation: guidelines for developing a customized suite of innovation metrics. *Strategy & Leadership*, *33*(1), 37–45.
Neely, A., Kennerley, M., & Adams, C. (2007). Performance measurement frameworks: a review. In A. Neely (Ed.), *Business performance measurement: unifying theory and integrating practice* (pp. 143–162), Cambridge.
Ojanen, V., & Vuola, O. (2006). Coping with the multiple dimensions of R&D performance analysis. *International Journal of Technology Management*, *33*(2/3), 279–290.
Page, A. L., & Schirr, G. R. (2008). Growth and development of a body of knowledge: 16 years of new product development research. *Journal of Product Innovation Management*, *25*(3), 233–248.
Pelz, D. C. (1978). Some expanded perspectives on use of social science in public policy. In M. Yinger & S. Cutler (Eds.), *Major social issues: a multidisciplinary view* (pp. 346–357), New York.
Pillai, A. S., Joshi, A., & Rao, K. S. (2002). Performance measurement of R&D projects in a multi-project, concurrent engineering environment. *International Journal of Project Management*, *20*(2), 165–177.
Rich, R. F. (1977). Use of social science information by federal bureaucrats: knowledge for actions vs. knowledge for understanding. In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 199–211), Lexington.
Roussel, P. A., Saad, K. N., & Erickson, T. J. (1991). *Third generation R&D: managing the link to corporate strategy*. Boston.
Sandstrom, J., & 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.
Sandt, J. (2004), *Management mit Kennzahlen und Kennzahlensystemen*. Wiesbaden.
Schäffer, U. (2007). *Management accounting & control scales handbook*. Wiesbaden.
Schäffer, U., & Steiners, D. (2004). Zur Nutzung von Controllinginformationen. *Zeitschrift für Planung & Unternehmenssteuerung*, *15*, 377–404.
Schumann, P. A., & Ransley, D. L. (1995). Measuring R&D performance. *Research Technology Management*, *38*(3), 45.
Sharma, S., Mukherjee, S., Kumar, A., & Dillon, W. (2005). A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. *Journal of Business Research*, *58*(7), 935–943.
Simon, H. A., Guetzkow, H., Kozmetsky, G., & Tyndall, G. (1954). *Centralization vs. decentralization in organizing the controller's department: a research study and report*. New York, NY.
Simons, R. (1994). How new top managers use control systems as levers of strategic renewal. *Strategic Management Journal*, *15*(3), 169–189.
Souchon, A. L., & Diamantopoulos, A. (1996). A conceptual framework of export marketing information use: key issues and research propositions. *Journal of International Marketing*, *4*(3), 49–71.
Souchon, A. L., Diamantopoulos, A., Holzmuller, H. H., Axinn, C. N., Sinkula, J. M., Simmet, H., & Durden, G. R. (2003). Export information use: a five-country investigation of key determinants. *Journal of International Marketing*, *11*(3), 106–127.
Stivers, B. P., Covin, T. J., Hall, N. G., & Smalt, S. W. (1998). How nonfinancial performance measures are used. *Management Accounting*, *79*(8), 44–48.
Toften, K., & Olsen, S. O. (2004). The relationships among quality, cost, value, and use of export market information: an empirical study. *Journal of International Marketing*, *12*(2), 104–131.
Ullrich, M. J., & Tuttle, B. M. (2004). The effects of comprehensive information reporting systems and economic incentives on managers' time-planning decisions. *Behavioral Research in Accounting*, *16*, 89–105.
Utterback, J. A. (1994). *Mastering the dynamics of innovation*. Boston.
Van der Stede, W. A., Young, S. M., & Chen, C. X. (2005). Assessing the quality of evidence in empirical management accounting research: the case of survey studies. *Accounting, Organizations & Society*, *30*(7/8), 655–684.
Vandenbosch, B. (1999). An empirical analysis of the association between the use of executive support systems and perceived organizational competitiveness. *Accounting, Organizations & Society*, *24*(1), 77–92.
Werner, B. M., & Souder, W. E. (1997). Measuring R&D performance: state of the art. *Research Technology Management*, *40*(2), 34–42.
West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with no-normal variables: problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling* (pp. 56–75), Thousand Oaks, CA.
Wheelwright, S., & Clark, K. (1992). *Revolutionizing product development: quantum leaps in speed, efficiency, and quality*. New York.