

Bican, Peter M.; Brem, Alexander

Article — Published Version

Managing innovation performance: Results from an industry-spanning explorative study on R&D key measures

Creativity and Innovation Management

Provided in Cooperation with:

John Wiley & Sons

Suggested Citation: Bican, Peter M.; Brem, Alexander (2020) : Managing innovation performance: Results from an industry-spanning explorative study on R&D key measures, Creativity and Innovation Management, ISSN 1467-8691, Wiley, Hoboken, NJ, Vol. 29, Iss. 2, pp. 268-291, <https://doi.org/10.1111/caim.12370>

This Version is available at:

<https://hdl.handle.net/10419/230086>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>

REGULAR ARTICLE

Managing innovation performance: Results from an industry-spanning explorative study on R&D key measures

Peter M. Bican¹  | Alexander Brem^{2,3} 

¹Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Chair of Technology Management, Nuremberg, Germany

²University of Stuttgart, Institute of Entrepreneurship and Innovation, Stuttgart, Germany

³University of Southern Denmark, Mads Clausen Institute, Sønderborg, Denmark

Correspondence

Peter M. Bican, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Chair of Technology Management, Nuremberg, Germany.

Email: peter.bican@fau.de

Research on R&D performance measures applied in firms is still scarce. Based on the established “R&D laboratory as a system” thinking, systematically derive and identify R&D department level key performance measures. Through a mixed-method approach, grounded in (1) literature and (2) text analysis, 154 R&D performance measures were developed. Amongst those, an (3) online expert survey, as well as (4) three independent focus group workshops with >40 industry experts from more than ten industries identified and validated ten key R&D performance measures. All industry experts involved are members of an innovation network, additionally accounting for innovation network effects. In contrast to earlier research, some of the measures like *degree of anticipation of internal customer needs* were perceived both by the survey respondents and the focus groups as key measures, indicating that behavioral measures should not be excluded per se. However, the importance of external validity of R&D performance or indicators to measure performance in relation to activities outside the R&D department were not confirmed. Hence, we partly confirm the relevance of the original “R&D Lab” measures, contributing a more granular level, thereby drawing implications for future research and practice.

KEYWORDS

innovation management, KPI, performance indicators, performance measurement, R&D management, R&D performance, innovation measures

1 | INTRODUCTION: WHY R&D PERFORMANCE MATTERS

Resource allocation in firms is a key managerial task. Dynamic capabilities were identified to play an important role in performing such tasks (Teece, Pisano, & Shuen, 1997). Therefore, firms need to persistently adopt internal and external capabilities. How firms handle these capabilities influences, among others, the success of their innovation activities, which is usually quantified by performance measures (Teece, 2007). Central to performance in innovation management are research and development (R&D) departments, as these are at the

heart of firms' innovation outputs, at least in high-tech environments. A common challenge in this context is to measure the almost unmeasurable, namely the performance (Chiesa & Frattini, 2007; Tkotz, Munck, & Wald, 2018). Reluctance of the staff to cooperate, mostly because they fear permanent monitoring and self-justification, often impedes these efforts (Brown & Svenson, 1998). Despite these obstacles, financially viable firms seem to be more successful in mastering the challenge of measuring R&D performance (Cooper & Kleinschmidt, 1995; Griffin, 1997).

In addition to this internal challenge, collaboration is a way to increase firms' efficiency (Lee, Kim, & Choi, 2019). External partners

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. Creativity and Innovation Management published by John Wiley & Sons Ltd

can help to source knowledge and thus spur innovation in times of open innovation (Bican, Guderian, & Ringbeck, 2017; Chesbrough, 2003; Enkel, Gassmann, & Chesbrough, 2009), coopetition (Sellitto & Luchese, 2018), and evaluation (Hienert & Riari, 2015). The exchange of knowledge and ideas is also fostered by innovation networks, which continue to grow in number and importance and constitute a form of open innovation (Kale & Singh, 2009). Performance measures would be necessary here, even though they are not easy to implement (e.g. Brown & Svenson, 1998; Henttonen, Ojanen, & Puumalainen, 2016).

Surprisingly, research on applicable R&D performance measures is still scarce, even though there is a high demand from the industry (Birchall, Chanaron, Tovstiga, & Hillenbrand, 2011; Dziallas & Blind, 2019; Larsen & Lindquist, 2016). Most prior research applies only a limited number of measures to assess the effect of innovation on R&D performance, depending on the focal point of the respective study, like the number or quality of patents (in terms of citation counts), cost and resource allocation, or external cooperation (e.g. Beers & Zand, 2014; Detzen, Verbeeten, Gamm, & Möller, 2018; Khanna, Guler, & Nerkar, 2016). Innovation measures, which are difficult to originate or to practically link to R&D activities, like return on investment or time-to-market, are commonly used metrics (Englund & Ludvigsen, 2015; Kristiansen & Ritala, 2018; Larsen & Lindquist, 2016). Research distinguishing between innovation performance and R&D performance measures is scant and often focuses on an abstract concept or introduces only a limited number of measures (e.g. Brown & Svenson, 1998; Dziallas & Blind, 2019; Škec, Cash, & Štorga, 2017; Werner & Souder, 1997).

To assess how the performance of R&D activities can be measured, our analysis of performance measures focuses on the level of R&D activities within the R&D department only (unlike, for example, the Frascati Manual or the Oslo Manual), which includes the inputs, the processing system, and outputs (Brown & Svenson, 1998). The analysis of these performance measures is based on the concept of the “R&D laboratory as a system” by Brown and Svenson (1998), whose findings are challenged by our results.

This induces our research question of analyzing R&D performance measures that are used in firms, as well as their applicability to other firms. We contribute the following: Through a mixed-method, grounded in prior literature of innovation and R&D measurement and evaluation systems and text analysis, 154 R&D performance measures are developed and further condensed to 81 performance measures. These measures form a unique base to stimulate future research in performance measures and innovation network performance effects.

Additionally, through a descriptive online expert survey, as well as three independent focus group workshops with more than 40 industry experts from more than ten industries, we contribute a selection of the most relevant key measures. The industry experts identified and validated ten key R&D performance measures individually and at innovation network level. All industry experts involved are members of an innovation network, additionally accounting for innovation network effects. The obtained results confirm the general relevance of the original “R&D laboratory as a system” measures (Brown &

Svenson, 1998). However, they also indicate that a more nuanced and granular assessment and framing of performance measures might be required.

Based on the established “R&D laboratory as a system” thinking (Brown & Svenson, 1998), this paper systematically develops and derives R&D department level key performance measures through a mixed-method approach. Ten key R&D performance measures were identified and validated. Thereby, the relevance of the original “R&D Lab” measures were partly confirmed, drawing both implications for future research and practice.

2 | THEORETICAL BACKGROUND: WHAT DO WE KNOW ABOUT R&D PERFORMANCE MEASUREMENT?

Even though some literature on performance measurement can be found, there are only a few studies that focus on innovation performance measurement in connection with R&D management. In order to derive key measures of R&D performance from prior literature, we apply the definition of R&D by Wheelwright and Clark (1992, p. 75): “Research and development is the creation of the know-how and know-why of new materials and technologies that eventually translate into commercial development”. We furthermore rely on the established definition of performance, performance measurement, and performance measure by Neely, Mills, Platts, Gregory, and Richards (1996) (see, e.g., in Henttonen et al., 2016). Neely et al. (1996, p. 424) define performance as “the efficiency and effectiveness of action”, performance measurement as “the process of quantifying the efficiency and effectiveness of action”, and a performance measure as “a metric used to quantify the efficiency and/or effectiveness of action”.

Brown and Svenson (1998) designed a system for R&D measurement, the so-called “R&D laboratory as a system” (see Figure 1), which was identified to be a key concept that structures the link between R&D and performance measurement. The R&D laboratory gives researchers the opportunity to differentiate between various stages that are referred to as systems. They propose to look at R&D evaluation and measurement through the lenses of the R&D Lab as a system to understand success and failure of R&D measurements.

This R&D laboratory thinking spans five individual systems, namely (1) inputs, (2) the R&D Lab, (3) outputs, (4) the receiving system, and (5) outcomes. Following Brown and Svenson (1998, pp. 105–106), these stages are briefly described in the following:

- Inputs: “Inputs are the raw materials or stimuli a system receives and processes.”
- R&D Lab: “The Process System is the R&D Lab itself, which turns the inputs into outputs by writing proposals, conducting research, testing hypotheses, reporting results, and so on.”
- Outputs: “Typical outputs include patents, new products, new processes, [awards, presentations, (scientific)] publications, or simply facts, principles, or knowledge that were unknown before.” The

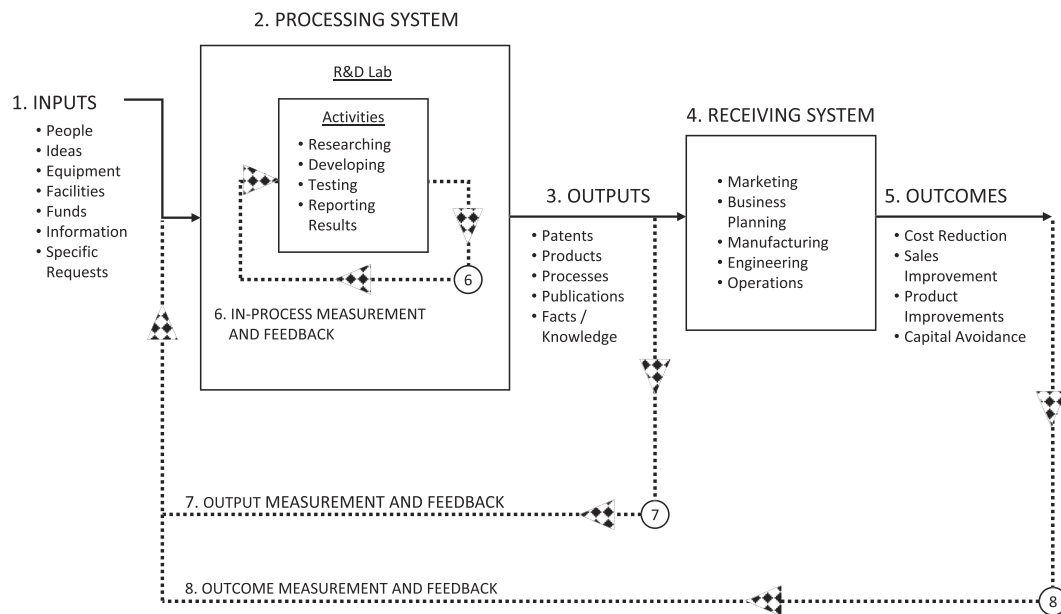


FIGURE 1 R&D laboratory as a system. Source: Brown and Svenson (1998, p. 106)

amount of these outputs is proposed as a measure of R&D performance.

- Receiving system: The various consumers of the R&D outputs are comprised, covering recipients like operations, marketing, and others.
- Outcomes: "Outcomes are the accomplishments that have value for the organization. They are produced when the receiving system accomplishes something with the outputs". Measures include sales volume, customer feedback, market share, or capital avoidance.

Brown and Svenson (1998) criticize that firms rely too much on internal process measurements and overly focus on behavior, which encourages the measurement of questionable outputs like the number of research proposals or published papers. They differentiate between measures of activity (i.e. what people do at work) versus measurement of accomplishments (i.e. what people achieve at work). Firms should focus on the latter to enhance the quality of the evaluation process, since behavior or activities are hard to measure (Brown & Svenson, 1998). If measurement focuses on activities, people care more about how they behave at work than what they achieve (Brown & Svenson, 1998). Moreover, these measures are too complicated or subjective to reliably evaluate R&D performance. To overcome these obstacles to measurement, Brown and Svenson (1998) present six characteristics of a successful measurement system:

- "Focus on external vs. internal measurement",
- "Focus on measuring outcomes and outputs, not behavior",
- "Measure only valuable accomplishments/outputs",
- "Make the measurement system simple",
- "Make the measurement system objective",
- "Separate R&D evaluation".

To apply the "R&D laboratory as a system" to various industries and incorporate additional theoretical streams, we further integrate findings from related R&D performance measurement literature, like R&D organizational literature that is mainly focused on measuring innovation activities within firms (Coccia, 2001; Henttonen et al., 2016; Kerssens-van Drongelen & Bilderbeek, 1999). For this purpose, we included classifications from the production system of the research bodies and balanced scorecard for an R&D department (Bremser & Barsky, 2004; Kerssens-van Drongelen & Bilderbeek, 1999; Kerssens-van Drongelen & Cooke, 1997).

In this vein, the focus of the current paper lies on project performance measures and not on organizational performance measures (Wang, Lin, & Huang, 2010), since innovation performance within the R&D department is measured at the R&D department level. Contrary to the definition of performance management provided by Kerssens-van Drongelen and Bilderbeek (1999) as "the acquisition and analysis of information about the actual attainment of firm objectives and plans, and about factors that may influence this attainment", we focus on the R&D Lab in terms of inputs and outputs. Since R&D should be considered separately (Brown & Svenson, 1998), the performance of the R&D department differs from the innovation outcomes as the total sales or the number of new product introductions to the market (as, e.g., in Lakiza & Deschamps, 2018). This follows our understanding of R&D as defined by Wheelwright and Clark (1992) who thought of R&D activities as creative activities before eventual commercial development, which is product marketing in its essence.

Asserting that some firms do not measure R&D performance at all, Kerssens-van Drongelen and Bilderbeek (1999) trace this reluctance to factors such as the degree to which R&D is connected to other departments like marketing or manufacturing. This makes it even more necessary to view the R&D Lab as a system, to ensure an

individual consideration of the different stages like inputs and outputs (Brown & Svenson, 1998).

Nevertheless, Kerssens-van Drongelen and Bilderbeek (1999) detected that around 80 percent of firms measure R&D performance. According to them, measures should also be adaptable to the respective type of research: (1) basic research, (2) applied research, and (3) development (Kerssens-van Drongelen & Bilderbeek, 1999). This follows Brown and Svenson (1998), who separate product development from research and technology management. Furthermore, Kerssens-van Drongelen and Bilderbeek (1999) compared measures versus measurement methods like benchmarking or organizational auditing in order to identify highly and robustly effective measurement procedures, especially in view of the customer. This adds an external element of performance evaluation to the R&D Lab discussion that has already been recognized before (see Brown & Svenson, 1998). However, the particular importance of external validity is further stressed and extended by complementing it with the external network and development measures of R&D performance measurement. Contrary to Brown and Svenson (1998), Kerssens-van Drongelen and Bilderbeek (1999) do not find the measurement of ideas to be of particular interest for evaluating R&D performance. These findings also hold true for measurements related to creativity.

Intangible assets, as main drivers of firm success (Conley, Bican, & Ernst, 2013), motivated Baglieri, Chiesa, Grando, and Manzini (2001) to carry out research on how to evaluate these assets. Performance measurement of technological assets is used as a starting point in this process. For example, outcomes of R&D activities might differ vastly from the R&D activities themselves, since the outcomes entail both the efforts within the R&D department as well as work from other departments, and often materialize only years after the underlying R&D project's completion (Baglieri et al., 2001). Traditionally, firms have been relying on input instead of output variables for measuring R&D performance on the individual level, which calls into question the assumptions made above. Building on Wheelwright and Clark (1992), Baglieri et al. (2001) define two phases of R&D activities: generative and transitionary, i.e. technological progress or knowledge transfer as outputs. This also covers the capacity to informally transfer knowledge, which is deemed important for scientific activities and is hence often underestimated (Azzone & Maccarrone, 1997; Coccia, 2001).

These measurement difficulties are also labeled as uncertainties, which are associated with risk in R&D success (Baglieri et al., 2001; Wang et al., 2010). To handle these, an R&D risk-framework is applied in the pharmaceutical context that does not differentiate between outputs and outcomes. Several R&D performance measures at the R&D department level are proposed, like the *employee retention rate*, *strategic skills coverage ratio*, i.e. *employees having the right strategic skill set to meet organizational needs*, or *percentage of collaborative projects with third parties*, which also does not distinguish between the different stages of the R&D process (Wang et al., 2010). In this way, Wang et al. (2010) propose to measure performance in the dimensions of predictability, staff performance, data quality, lead time, and on-time schedule.

Following the description by Kerssens-van Drongelen and Bilderbeek (1999), p. 38): "[P]urpose of the measurement and objectives formulated for the subject of management", Henttonen et al. (2016) identified the importance of performance measurements, thereby confirming differences in the role of evaluators (as e.g. Blindenbach-Driessen, van Dalen, & van den Ende, 2010). However, as in previous studies (e.g. Lazzarotti, Manzini, & Mari, 2011), Henttonen et al. (2016) did not account for the difference between outputs and outcomes, as proposed by Brown and Svenson (1998).

Within this section, a short motivation into performance measures and topics which are related to measurement, encouraging a culture of measuring R&D performance in firms, has been presented. As becomes visible, performance measurement is dispersed over multiple fields. Therefore, prior research has focused on various aspects, frameworks, and interpretations of performance measures in different contexts. Deriving a set of concrete measures of R&D performance and assessing this set of measures to identify the key measures in evaluating R&D performance has not yet been a dominant focus of scholarly attention (Hienerth & Riar, 2015). Too many measures have been introduced with only little testing, which calls for more research to better understand measurement factors (Birchall et al., 2011). This induces our research question of analyzing R&D performance measures that are used in firms, as well as their applicability to other firms. Based on the R&D Lab framework and prior works from the above presented authors, a set of measures has been derived and composed that will be evaluated in the following sections.

3 | METHOD AND DATA

To assess the role of R&D performance measures and to answer our research question, a mixed-methods approach has been applied (Johnson & Onwuegbuzie, 2004). This approach is anchored in our research question, i.e. how the performance of R&D activities can be measured (Johnson & Onwuegbuzie, 2004). In doing so, R&D performance measures that are used in firms are analyzed, as well as their applicability to other firms. Key performance measures most relevant to firms were identified. Mixed-methods have the advantage of reducing bias stemming from one method alone, so that through different types of data collection and data analysis, a richer and contextual understanding of the subject being researched can be reached (Denzin, 1978; Gray, 2018; Hanson, Creswell, Clark, Petska, & Creswell, 2005; Johnson & Onwuegbuzie, 2004). In doing so, the insights revealed by the mixed-method approach add more value than insights that could have been revealed through a single method alone (McKim, 2017; Schoonenboom & Johnson, 2017). Therefore, both quantitative and qualitative elements guided the selection of key performance measures (see Figure 2).

3.1 | Literature review and analysis

The literature review enabled identification and a critical assessment of the relevant spectrum of prior research on performance

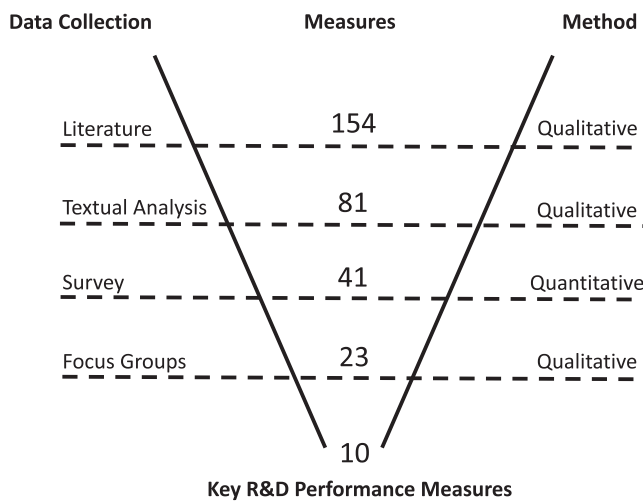


FIGURE 2 Overview of methodological steps and linked measure development

measurement, contrary to a systematic literature review (Branley, Seale, & Zacharias, 2018; Hart, 2018). Systematic literature reviews, both past and recently conducted, do not specifically focus or target R&D performance measures only (e.g. Dziallas & Blind, 2019; Werner & Souder, 1997). This lies also within the nature of systematic literature reviews, which aim to cover everything within their scope limitation (Gray, 2018). Often, systematic literature reviews are limited to certain journals and fields, hence neglecting adjacent publications or other types of publications like books or conference contributions (as, e.g., in Dziallas & Blind, 2019; Werner & Souder, 1997). Since performance measures, R&D, and innovation intersect various research streams like innovation management, accounting, or controlling, systematic literature reviews, with their narrow focus and clearly defined exclusion criteria, would not satisfy the research aim of identifying the broadest set of R&D measures from prior literature (e.g. Jesson, Matheson, & Lacey, 2011). Consequently, literature was queried on databases like Scopus and Google Scholar, employing a wide range of various search terms related to innovation, R&D, performance, and measurement, to identify the relevant literature for the measure identification, based on the concept of the “R&D laboratory as a system” by Brown and Svenson (1998).

A deep textual analysis revealed that measures from different papers overlapped in parts or were not applicable to this paper's R&D Lab focus. We analyzed each of the measures independently concerning their content, reach, and positioning within the R&D Lab (Brown & Svenson, 1998). This enabled us to condense the 154 measures identified from the literature review to 81 unique measures, centering around inputs, R&D Lab, and outputs. Triangulation was ensured through two independent assessments by both authors. In case of deviation, consensus was reached through discussion. Thereby, unclear or equivocal measures were edited to enhance the understanding of the measures' meaning.

3.2 | Participants of the descriptive survey and focus group workshops

The participants of both the descriptive survey and the focus groups are members of an innovation network, which goes beyond industry borders. Our findings profit from the specialty of the networks' organizational composition and its governing structure: Organized in the form of a German non-profit-organization called “Verein” (association), members are loosely connected (e.g. in the form of loose networking events and working groups). This creates a form of safe environment that is beneficial for research purposes, for example when conducting focus group workshops (Freeman, 2006).

Within this association, there are also several working groups. One of these groups is concerned with innovation performance measurements, as this seems to be a challenging task in many different firms. This working group forms the basis of our samples, both for the descriptive survey and the first and second working groups, whereby focus group data could be collected from various independent institutions that are active in different industries (as the composition of participants differed between the individual working groups).

3.3 | Descriptive survey

To surface facts and to test and assess the relevance of the identified measures, we applied a descriptive survey method (Gray, 2018). Contrary to the analytical survey, descriptive surveys are designed to measure certain characteristics, hence answer the “what” instead of the “why” (Gray, 2018). Data from an anonymous expert survey that included ten firms from different industries were collected via an online questionnaire. All respondents were experienced R&D or innovation managers. The descriptive survey was conducted anonymously to ensure respondents' confidentiality and openness. In this empirical setup, the respondents were asked to answer the following questions with yes or no for each of the 81 measures:

- Do you know this measure (question 1),
- Is this measure useful for your firm (question 2), and
- Is this measure applied in your firm (question 3).

Based on the overall selection of participants' answers to the above questions, a prioritization list of 41 measures was derived. These measures had been deemed as either useful (i.e. participants answered “yes” to question 2) or practically applied (i.e. participants answered “yes” to question 3). These results constituted the foundation for the identification of key performance measures. Additionally, as there might be differences between participants' answers in anonymous surveys to actual behavior in the real world (Black, 1993), the results from the descriptive survey

were further analyzed and challenged through three consecutive focus group workshops.

3.4 | Focus group workshops

Focus groups are appropriate for this purpose and to further develop the afore-derived measures: They are advantageous in shedding light on the different views of participants, independent of the respective topic (Gray, 2018; Stewart & Shamdasani, 2015). Furthermore, they provide opportunities for clarification of responses and in assessing and valuing opinions and positions (Freeman, 2006; Seal, Bogart, & Ehrhardt, 1998; Stewart & Shamdasani, 2015). As our research purpose is to analyze R&D performance measures that are used in firms, as well as their applicability to other firms, and to identify key performance measures most relevant to firms, focus group workshops were applied threefold, whereby each distinct workshop extends the findings of the previous workshop (see Figure 3).

In designing, organizing, and executing the focus group workshops, we followed the approach outlined by Stewart and Shamdasani (2015): Problem identification, sampling frame/recruiting, moderator, facilities, interview questions, group execution, data recording, and data analysis. The focus group workshops were conducted between November 2017 and October 2018 to use the knowledge of industry experts on R&D performance measures. For this purpose, participants of the focus groups came from different firms and industries. All participants are experienced R&D or innovation managers, with firm tenures of up to 25 years. Hence, an across-industry view on these measures could be gathered. Participants were introduced to the methodology of our analysis, the distinction between inputs, R&D Lab, and output stages, as well as the descriptive survey results.

Participants in the first focus group workshop were asked to further condense the list of 41 measures and to choose their top ten and flop ten measures. Participants were to do so without knowing which stage each measure belonged to in order to avoid early-stage selection bias. The assessment of the top ten and flop ten was conducted in three randomly assigned groups with the following group characteristics: Group 1 consisted of four managers, who are active in technology and innovation and predominantly work in academic or consulting

environments. Group 2 had nearly the same setup, with all four members active in innovation management roles in the automotive or related industries. Group 3 was rather heterogeneous, with five group members from innovation management and controlling, who are employed in industries like mechanical engineering, automotive, insurance, and mechatronics.

The participants were given the instruction to choose the ten measures, which they would refer to as key R&D performance measures. This enabled structured sampling to ensure the generalizability of the results. However, it still remains an exploratory analysis, since the sample size is limited. However, the small sample size also allowed for a more detailed analysis of the participants' evaluation and understanding of measures. This contributed additional value, e.g. in the form of insights on why measures are not understood or used.

Contrary to the first focus group workshop, the setting was modified for the second focus group workshop by combining elements from both the descriptive survey and the first focus-group workshop: During the preparation, the 41 measures were handed out upfront to all participants. These instructions included additional explanations for each measure, also covering the related stage each measure was attributed to. In this way, the results of this focus group could be further scrutinized and compared to the first focus group workshop, revealing insights into measure adoption and understanding. Participants were asked the same three questions and, in addition, to rank the performance measures according to their importance and thus identifying an additional set of key performance measures.

The third focus group varied in composition and scope: To extend the reach beyond the respective network's working group member firms, a third, final focus group workshop was organized. This workshop with 20 participants was expanded to the innovation network level to validate and discuss the previously identified key R&D performance measures with participants different from before (i.e. in focus groups 1 and 2). Like in the descriptive survey, participants were asked the three questions for each of the key R&D performance measures, to find out if they had prior experience in working with innovation measures. In this way, the research scope could be further extended from the innovation network's working groups level to the full innovation network, enhancing innovation at the innovation network level.

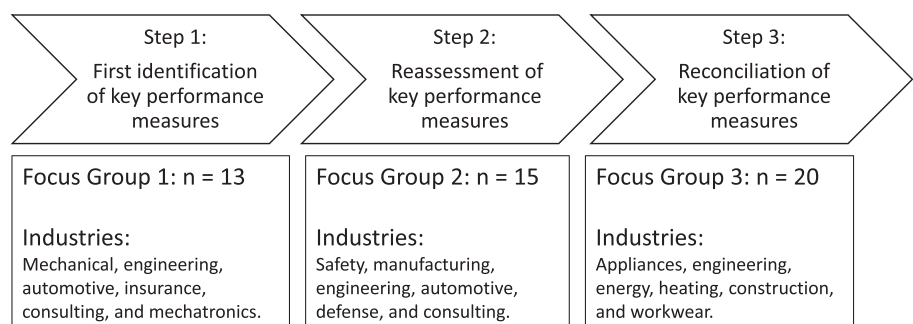


FIGURE 3 Overview of focus group workshops

4 | RESULTS

4.1 | Literature review and descriptive survey

The literature review revealed a fragmented and diverse field on performance measures, spanning multiple research areas and levels of analysis. For example, Birchall et al. (2011) split their analysis of performance measures as to their scope and nature, with five measurement scales and findings on 27 indicators of performance: (1) future focus, (2) market impact, (3) capabilities and image, (4) process, and (5) sustainability and overall effectiveness. In contrast, Chiesa, Frattini, Lazzarotti, and Manzini (2007) and Lazzarotti et al. (2011) propose to measure R&D performance with a formal model by applying quantitative indicators from five perspectives of performance: the (1) financial perspective, (2) customer perspective, (3) innovation and learning perspective, (4) internal business perspective, (5) alliances and networks perspective. By basing their measurements on soft systems, they aim “to analyze the measurability of performance in R&D processes of a company as the system under measurement” (Lazzarotti et al., 2011).

Various analyses also use a diverse set of methods: Henttonen et al. (2016), for example, measured project success by querying 169 respondents of the same firm on performance measures for various projects and project types. Amongst their five highest selected measures, the level of innovativeness, effort and commitment to objectives, time to market, customer satisfaction, and degree of multifunctional cooperation, at least two can be characterized as outcomes (i.e. time to market and customer satisfaction), impeding a derivation of R&D performance measures only. Focusing on R&D effectiveness and efficiency, Chiesa and Masella (1996) propose to distinguish between measures that can be applied before or after project exploitation, hence completion. Measures comprise technical success and measures of efficiency of the R&D department. Coccia (2001) extends this thinking to the research laboratories' evaluation methodology to obtain a scoring for financial, scientific, and technological aspects through simple indices. Often, proposed indicators are applied to one case only and do not distinguish between outputs and outcomes, as suggested by Brown and Svenson (1998). As shown through the above examples (see a summary of the most relevant papers identified in Table 1), different measurement perspectives are proposed, for example with a focus on design, like education and competency development or learning capability (Škec et al., 2017).

Eventually, the literature review revealed 154 measures that were identified and collected from the reviewed papers (see Appendix). These were further condensed to 81 measures. The descriptive survey indicated that 55 measures (68 percent) of these previously condensed 81 measures were known, 40 measures (49 percent) were deemed useful, and 24 measures (30 percent) were applied within the participating firms. The sample of 41 measures identified through the descriptive survey consisted of 11 input, four R&D Lab, and 26 output measures. The results of the descriptive survey are summarized in Table 2.

Compared to the ratio of measures between input, R&D Lab, and output stage between the condensed literature review analysis results

and the descriptive survey, the ratio of R&D Lab classified measures further decreased (from 18 to 10 percent) in favor of output classified measures (from 55 to 63 percent). The ratio of input measures stayed consistent at 27 percent. Table 3 presents an overview of the ratio of the three stages.

4.2 | Focus group workshop 1

The 41 measures identified through the descriptive survey were assessed in a focus group setting (see results in Table 2). The assessment of the three groups (range of score: zero to three) revealed the following key and flop measures. To focus on a few, selected measures only, each group had to limit their selection to 10 out of 41 measures. The groups selected the measures as shown in Table 4.

The following three measures were not selected:

- Number of designs produced (Outputs)
- Number of ideas or findings employed (Outputs)
- People involved, i.e. number of hours worked (Inputs)

The key selected measures (i.e. key performance measures), the flop selected measures, and measures that were treated equivocally (considered to be both flop and key measures) are presented in Table 5.

4.3 | Focus group workshop 2

To validate these findings and further analyze the importance of the performance weighting, the six key selected measures of the second focus group workshop resulted in two measures overlapping with the results from focus group workshop 1. Table 6 summarizes the key selected measures of both focus group workshops.

Focus group 2 also assessed the 41 measures with the questions if the measures were known, useful, and applied within their firm, mirroring the original descriptive survey approach to pre-select relevant performance measures.

4.4 | Focus group workshop 3

With focus group workshop 3, we aimed to analyze the previously selected key measures in terms of their innovation network level validity. During this third focus group workshop, the most important measures of both previous focus groups were jointly presented as a selection of the ten key performance measures. Answers confirmed the identified key measure selection, since all but one key measure were known to the focus group participants, except for *degree of anticipation of internal customer needs*. The key measures were furthermore deemed to be useful, with the exception of *hours spent on projects vs. total hours R&D*.

TABLE 1 Summary of key literature

| Title | Authors | Journal | Volume | Issue | Year | Pages | Main contribution |
|---|---|---|----------|-------|------|-----------|--|
| Measuring R&D Productivity | Mark G. Brown Raynold A. Svenson | Research Technology Management | 31 | 4 | 1998 | 105–110 | R&D Lab as a system |
| Searching for an effective measure of R&D performance | Vittorio Chiesa Christina Masella | Management Decision | 34 | 7 | 1996 | 49–57 | R&D effectiveness and efficiency |
| R&D Performance Measurement: More than choosing a set of metrics | Inge c. Kerstens-van Drongelen Jan Bilderbeek | R&D Management | 29 | 1 | 1999 | 35–46 | Highly and robust effective measurement procedures |
| A basic model for evaluating R&D performance | Mario Coccia | R&D Management | 31 | 4 | 2001 | 453–464 | Mathematical model of public laboratory like systems |
| Evaluating Intangible Assets: The Measurement of R&D Performance | Enzo Baglieri Vittorio Chiesa Alberto Grando Raffaella Manzini | Research Division Working Paper No. 01/49 | | | 2001 | | Firm's value and R&D shareholder value creation |
| A performance-oriented risk management framework | Juite Wang Willie Lin Yu-Hsiang Huang | Technovation | 30 | 11/12 | 2010 | 601–611 | Aligning measurement to mitigate R&D risks for R&D success |
| Innovation performance measurement: Current practices, issues and management challenges | David Birchall Jean-Jacques Chanaron George Tovstiga Carola Hillenbrand | International Journal of Technology Management | 56 | 1 | 2011 | 1–20 | Five measurement scales for innovation performance |
| A model for R&D performance measurement | Valentina Lazzarotti Raffaella Manzini Luca Mari | International Journal of Production Economics | 134 | 1 | 2011 | 212–223 | A formal model for measuring R&D performance |
| How should firms evaluate success in university–industry alliances? | Markus Perkmann Andy Neely Kathryn Walsh | R&D Management | 41 | 2 | 2011 | 202–216 | Success map to distinguish process stages |
| Searching for appropriate performance measures for innovation and development projects | Kaisa Henttonen Ville Ojanen Kaisu Puumalainen | R&D Management | 46 | 5 | 2016 | 914–927 | Dimensions of measuring performance |
| A dynamic approach to real-time performance measurement in design projects | Stanko Škec Phillip Cash Mario Štorga | Journal of Engineering Design | 28 | 4 | 2017 | 255–286 | Fully realized performance measurement approach |
| Exploring data envelopment analysis for measuring collaborated innovation efficiency of small and medium-sized enterprises in Korea | Jiyoung Lee Chulyeon Kim Gyunghyun Choi | European Journal of Operational Research | In Press | | 2018 | | Efficiency of collaboration |
| Formal controls and team adaptability in new product development projects | Nina Detzen Frank H.M. Verbeeten Nils Gamm | Management Decision | 56 | 7 | 2018 | 1541–1558 | Target rigidity and performance goals |

(Continues)

TABLE 1 (Continued)

| Title | Authors | Journal | Volume | Issue | Year | Pages | Main contribution |
|---|--|---------------------------------|--------|-------|------|---------|--|
| Measuring radical innovation project success: typical metrics don't work | Klaus Möller Jimmi N. Kristiansen Paavo Ritala | Journal of Business Strategy | 39 | 4 | 2018 | 34–41 | KPIs for radical innovation projects |
| Evaluating firms' R&D performance using best worst method | Negin Salimi Jafar Rezaei | Evaluation and Program Planning | 66 | 2 | 2018 | 147–155 | Multi-criteria framework for firms' R&D performance evaluation |
| Innovation indicators throughout the innovation process: An extensive literature analysis | Marisa Dziallas Knut Blind | Technovation | 80/81 | 2/3 | 2019 | 3–29 | Process over product indicators |

5 | DISCUSSION AND IMPLICATIONS

With this research, we aimed to explore R&D performance measures used within firms, and to elevate these to key R&D performance measures, which are also suitable for quantifying performance outside firms' boundaries, i.e. in innovation networks. Furthermore, we reassessed the original R&D laboratory measures by Brown and Svenson (1998) in the light of our findings. Brown and Svenson's (1998) original R&D Lab included 17 measures in their core model (Figure 1), and 20 additional variables throughout the paper. Table 7 aims to align the measures from this paper with the selected key performance measures, and summarize them in three performance measurement categories that are in line with the Brown and Svenson (1998) paper.

As can be seen, none of the identified key performance measures were directly included in the core R&D Lab framework (see Brown & Svenson, 1998). Some, like *% of budget spent internally on applied research*, could be reconciled under the broader terms of *Researching*, *Developing*, or *Funds invested* within the original core framework. Our results are also not influenced by project success or R&D effort outcomes, since we did neither distinguish between successful or failed projects, nor mix output effects with outcome effects, focusing only on the R&D department. The mixed-method approach includes three distinct steps: (1) derivation of R&D performance measures from prior literature and textual analysis, and (2) assessment and analysis through a descriptive expert survey. The results from the descriptive expert survey were further discussed in three distinct focus group setups with industry experts and members of an innovation network. Hence, these results are useful for individual firms, but also for innovation networks.

In line with the observations of Kerssens-van Drongelen and Bilderbeek (1999), who noticed that around 80 percent of firms measure R&D performance in some way, our descriptive expert survey revealed that around 68 percent of the 81 proposed measures were known by our respondents, whereas only 40 percent were deemed useful and 30 percent applied in practice. These results must be scrutinized with regards to the discussions within our focus groups. Most participants claimed to struggle with finding any R&D measures within their firms or were wondering how R&D could be measured at all. This is surprising both in the light of our results and prior literature (e.g. Griffin, 1997; Kerssens-van Drongelen & Bilderbeek, 1999), but might be due to the fact that people who are active in R&D are not aware of performance activities actually being measured.

The evaluation of ideas does not seem to be especially important when assessing related measures. Even though the number of ideas is an important input factor and the *number of ideas or findings employed* scored high on all dimension and was deemed as useful in the expert descriptive survey, this measure was not selected at all within the focus groups, in line with the related measure of *percentage of project evaluation ideas applied in new projects* (only known by survey respondents and therefore not included in the focus group sample). This contradicts a core measure of Brown and Svenson (1998) and confirms the findings of Kerssens-van Drongelen and Bilderbeek (1999) who

TABLE 2 Descriptive survey results

| Measure | Level | Known | Useful | Applied | Selected | Measure | Level | Known | Useful | Applied | Selected |
|--|---------|-------|--------|---------|----------|--|---------|-------|--------|---------|----------|
| The ratio of technical progress to time | Outputs | YES | YES | YES | YES | Percentage of collaborative projects with third parties | R&D Lab | YES | | | |
| Technology leadership in terms of number of patents developed | Outputs | YES | YES | YES | YES | Number of technology transfer plans | Outputs | YES | | | |
| Sum of planned or revised project durations | Outputs | YES | YES | YES | YES | Number of research proposals | R&D Lab | YES | | | |
| Project progress and projects completed | R&D Lab | YES | YES | YES | YES | Number of reports, i.e. quality of documentation | Outputs | YES | | | |
| Project feasibility with milestones/gates | Outputs | YES | YES | YES | YES | Number of publications (books and paper) | Outputs | YES | | | |
| Product development cycle time | R&D Lab | YES | YES | YES | YES | Number of personnel in training | Inputs | YES | | | |
| Planning accuracy, i.e. % of agreed milestones and/or objectives met | Outputs | YES | YES | YES | YES | Number of awards won | Outputs | YES | | | |
| People involved, i.e. number of hours worked | Inputs | YES | YES | YES | YES | Number of alliances dedicated to technological innovation | R&D Lab | YES | | | |
| Number of redesigns and average time of redesigns | Outputs | YES | YES | YES | YES | Network built externally | Outputs | YES | | | |
| Number of products designed | Outputs | YES | YES | YES | YES | Leverage factor, i.e. factor by which funding is increased via public grants, etc. | Inputs | YES | | | |
| Number of parts produced | Outputs | YES | YES | YES | YES | Human capital indicators | Inputs | YES | | | |
| Number of new projects initiated | Outputs | YES | YES | YES | YES | Committee work performed | Inputs | YES | | | |
| Number of ideas or findings employed | Outputs | YES | YES | YES | YES | % of project evaluation ideas applied in new projects | Outputs | YES | | | |
| Number of funds invested | Inputs | YES | YES | YES | YES | % of budget spent externally on basic research | Inputs | YES | | | |
| Number of designs produced | Outputs | YES | YES | YES | YES | The ratio of technical progress to costs | Outputs | YES | YES | | |
| Hours spent on projects vs. total hours R&D | Inputs | YES | YES | YES | YES | The ratio of cost to time and to time per technical progress unit | Outputs | YES | YES | | |
| Design performance, i.e. as to manufacturing cost, manufacturability, or testability | Outputs | YES | YES | YES | YES | Rate of re-use of standard designs, i.e. as proven technology | Outputs | YES | YES | | |
| Degree of project attractiveness | Outputs | YES | YES | YES | YES | Degree of anticipation of internal customer needs | Outputs | YES | YES | | |
| Cost efficiency, i.e. keeping within budget | Outputs | YES | YES | YES | YES | Time dedicated to the analysis of reasons for failure of previous projects | Inputs | | | | |
| Average cost of each completed project | Outputs | YES | YES | YES | YES | The ratio of time to technical progress, i.e. the time employed to improve technical progress unit | Outputs | | | | |

(Continues)

TABLE 2 (Continued)

| Measure | Level | Known | Useful | Applied | Selected | Measure | Level | Known | Useful | Applied | Selected |
|---|---------|-------|--------|---------|----------|---|---------|-------|--------|---------|----------|
| Amount of facilities like laboratories, libraries, and equipment employed | Inputs | YES | YES | YES | YES | The ratio of technical progress to time, i.e. the technical progress per time unit (month, year, etc.) | Outputs | | | | |
| % of budget spent internally on applied research | Inputs | YES | YES | YES | YES | The number of project parts jointly carried out with other projects internally | Outputs | | | | |
| % of budget spent externally on applied research | Inputs | YES | YES | YES | YES | The number of project parts jointly carried out with other projects externally | Outputs | | | | |
| Transfer rate of new knowledge and technology into product development | Outputs | YES | YES | | YES | The degree of integration with the marketing function | R&D Lab | | | | |
| Quality, i.e. number of times reworked | Outputs | YES | YES | | YES | The degree of integration with the manufacturing function | R&D Lab | | | | |
| Quality of project management | R&D Lab | YES | YES | | YES | Strategic skills coverage ratio, i.e. employees having the right strategic skill set to meet organizational needs | Inputs | | | | |
| Percent of new technology content in new products | Outputs | YES | YES | | YES | Specific requests demanded | Inputs | | | | |
| Number of other legal protection methods besides patents | Outputs | YES | YES | | YES | Percentage of people having a pertinent degree (with respect to all the people devoted to technological innovation) | Inputs | | | | |
| Innovation level and degree of creativity | Inputs | YES | YES | | YES | Outside data integration, i.e. from other departments than R&D | Inputs | | | | |
| Evaluation of scope of patent applications | R&D Lab | YES | YES | | YES | Number of solution concepts, i.e. to particular problems without specifying exact technical recipes | Outputs | | | | |
| Degree of flexibility of resources employed | Inputs | YES | YES | | YES | Number of presentations made | Outputs | | | | |
| Annual spending for market investigations aimed at generating technological innovation | Inputs | YES | YES | | YES | Number of employees dedicated to external relationships in R&D | R&D Lab | | | | |
| % of projects technically successful | Outputs | YES | YES | | YES | Knowledge and information generated | R&D Lab | | | | |
| % of projects in which customers are operatively involved (training, testing, or problem solving) | Outputs | YES | YES | | YES | Facts and principles generated | Outputs | | | | |
| % of projects abandoned after a certain degree of completion | Outputs | YES | YES | | YES | Dependence on centralized funding, i.e. capacity for self-financing | Inputs | | | | |
| % of budget spent internally on basic research | Inputs | YES | YES | | YES | Degree of anticipation of external customer needs | Outputs | | | | |

TABLE 2 (Continued)

| Measure | Level | Known | Useful | Applied | Selected | Measure | Level | Known | Useful | Applied | Selected |
|---|---------|-------|--------|---------|----------|--|---------|-------|--------|---------|----------|
| Employee retention rate | Inputs | YES | YES | YES | YES | Cost methods like reproduction cost, updated cost, or cost of loss | Outputs | | | | |
| Staff learning metrics like number of teaching courses held | Outputs | YES | | | | Competitive value of technology innovations and research data | Outputs | | | | |
| Ratio of net present cash flow to development cost | Outcome | YES | | | | Combine data with internal conditions | R&D Lab | | | | |
| Ratio between researching and development activities | R&D Lab | YES | | | | Combine data with external developments | R&D Lab | | | | |
| Project efficiency in terms of number of duplications | Outputs | YES | | | | | | | | | |

stated that the evaluation of ideas is ancillary. However, the factor of *ideas* might be reflected in the more granular key measure of *innovation level and degree of creativity*.

Besides the measure *hours spent on project vs. total hours R&D*, no *people* (as in the original R&D Lab) or human resource focused measure was selected as a key performance measure. The *employee retention rate* was the sole measurement in the descriptive survey sample that was applied in firms though not deemed useful, which reinforces the notion that human resource factors might be deemed of lower importance in R&D performance measurement. Other measures like the *percentage of people having a pertinent degree or employees' strategic skills coverage ratio* were overwhelmingly unknown. This contradicts prior research (e.g. Perkmann, Neely, & Walsh, 2011; Wang et al., 2010) and is further reflected in the focus groups' assessment, with *people involved* not being selected at all and the measure *employee retention rate* being the only measure that was discarded by all three groups (focus group workshop 1) in unison. It seems as if human resource centric measures might be suitable in theory, but lack acceptance in practice which might be due to cases of staff reluctance to cooperate in measuring R&D performance, caused by fear of, for example, constant surveillance and monitoring (Brown & Svenson, 1998). This resembles discussions within the focus groups concerning operators or evaluators of the performance measures who leave the participants at odds with potential solutions.

As discussed earlier, the focus groups revealed a common understanding that no suitable measures seem to exist, at least when engaging in discussions before analyzing the descriptive survey results. However, when the survey results were discussed amongst the focus group participants, 68 percent of all measures were known, more than half of these applied and almost half deemed as useful. Furthermore, of the descriptive survey's original selection of 41 measures, only three were not once selected by focus group 1 as a key or flop measure, displaying the diverse views on a wide spread of potentially applicable measures as assessed by the groups (e.g. 23 of 41 measures were identified as the ten key measures by the groups). This is also confirmed with the results of the other focus groups, with 70 percent of all selected key measures being identified as known by focus group 3.

The importance of technology transfer has mainly been assumed in prior literature, with little validation, and is not present as a core measure in the original R&D Lab model (Brown & Svenson, 1998; Coccia, 2001). However, the *transfer rate of new knowledge and technology into product development* was assessed as known and useful in the descriptive survey and was identified as one of only two key R&D performance measures that were selected by both focus groups, contrary to *number of technology transfer plans*, which did not pass the descriptive survey stage (not assessed as applied or useful). This might have been the case, because the term "transfer plans" was too vague as a suitable measure or the content of this measure was already being reflected in the measure *transfer rate of new knowledge and technology into product development*. Interestingly, it is also one of two key measures that are both known and deemed useful by focus group 3, but not applied within the respective firms. This could also reflect the fact,

TABLE 3 Stage ratio per data sampling step

| Step | Literature | | Textual analysis | | Survey | | Focus Group 1 | | Focus Group 2 | | Focus Group 3 | |
|----------|------------|-----|------------------|-----|--------|-----|---------------|-----|---------------|-----|---------------|-----|
| Inputs | 22 | 14% | 22 | 27% | 11 | 27% | 5 | 22% | 1 | 17% | 3 | 50% |
| R&D Lab | 14 | 9% | 14 | 17% | 4 | 10% | 3 | 13% | 1 | 17% | 1 | 17% |
| Outputs | 74 | 48% | 45 | 56% | 26 | 63% | 15 | 65% | 4 | 66% | 2 | 33% |
| Outcomes | 44 | 29% | | | | | | | | | | |
| N | 154 | | 81 | | 41 | | 23 | | 6 | | 6 | |

as one participant stated, that “[w]e prioritized as to what we deemed useful” (Participant of focus group workshop 2, July 19, 2018).

While the number of identified input measures constantly matched around 30 percent of the measures identified in the textual analysis and the descriptive survey, this number dropped after the first focus group assessment to almost the same low levels as after the literature analysis (around 15 percent). This is in line with findings concerning the measure *ideas*, as discussed above (Kerssens-van Drongelen & Bilderbeek, 1999). However, when considering the key measure selection of focus group 2, the input measures almost match the outputs within the final ten key measures (40 percent vs. 50 percent of all key performance measures). It seems as if providing participants with information on the stage of measures and additional explanations per measure, enhances their understanding of the joint relevance of both input and output measures. Possibly, the traditional role of inputs is not overstated in practical application, which contradicts the assumptions by Baglieri et al. (2001).

6 | THEORETICAL IMPLICATIONS: EXTENSION OF THE R&D LAB MODEL

Our findings impact both research and practice of R&D performance measurement. Research has so far only focused on narrowly defined sets of measures, often without distinguishing between the different measurement dimensions. In addition, prior insights on R&D performance measures centered around outcome measures, which are difficult to apply in quantifying the innovation performance at the R&D department level. Research on deriving R&D key performance measures from an extensive set of R&D performance measures, is still scarce. Researchers thus profit from new insights into firms' application of R&D performance measures, a reassessment of previously identified frameworks like the R&D Lab by Brown and Svenson (1998), and the interplay with innovation network dynamics. Thereby, this paper contributes to the R&D organization literature, as well as innovation management and management accounting literature. As it becomes visible in the reassessment of the measures included both in the core R&D Lab model and introduced throughout the paper by Brown and Svenson (1998), previously described measures often lack detail and granularity, which complicates both their understanding and applicability (as, e.g., in Dziallas & Blind, 2019; Werner & Souder, 1997). Researchers can build on these findings and further assess and develop viable performance measures.

Furthermore, our findings indicate that a clear-cut classification of measures, e.g. to their relevant stage, proves challenging. Discussions in the focus group workshops revealed that measures could be applied both at the input or the output stage, even potentially spanning all stages of the R&D Lab. Taking this into account, the distinction between inputs and outputs could serve as a measurement of R&D lab activity instead. Reflecting on the “R&D laboratory as a system” thinking by Brown and Svenson (1998), as depicted in Figure 1, the role of inputs and outputs might make readjustment necessary. Instead of embracing input and output measures as separate dimensions, they might coalesce into the R&D lab dimension to enable the measuring of effects within the R&D lab dimension at any given point in time, depending on, for example, R&D project progress. Accounting for this at a more granular level would also enhance practical relevance and applicability. Figure 4 provides a view on how the respective R&D related parts of the original “R&D laboratory as a system” (Brown & Svenson, 1998) could be developed further in a tree-like structure (Menezes, Sellitto, Librelato, Borchardt, & Pereira, 2016; Sellitto, 2018) to reflect a less stringent categorization, including also the ten key performance measures identified.

In addition, similar to the perceived relevance of typical input measures, the supposed importance of typical output measures as listed by Brown and Svenson (1998) seemed to fade. Measures like the number of papers or books, awards won, or research proposals written, were identified through our expert descriptive survey as known, but neither rated as useful, nor applied within firms. Only measures related to completed projects, like received patents, designs produced, or products designed, were deemed to be useful and applied in our descriptive survey sample (e.g. *technology leadership in terms of number of patents developed, project progress and projects completed, number of products designed, number of designs produced*). Thus it can be concluded that measuring softer (i.e. less financially focused) outputs like scientific publications seems to be less important than hard facts like projects completed, as can also be seen in the selected key measures (as in Kerssens-van Drongelen & Bilderbeek, 1999; Chiesa & Masella, 1996).

Measures related to technical progress were included in various forms, stressing the importance of the current subject, just like in prior literature (Chiesa & Masella, 1996). Three of five measures included in the sample that entailed 81 performance measures, were positively assessed in the descriptive survey. However, the focus group ranked *the ratio of technical progress to costs* (useful in the survey assessment) and *the ratio of technical progress to time* (fully selected in the survey) as low, with only one group considering these measures to be

TABLE 4 Focus group measure selection results

| Measure | Level | KEY | | | FLOP | | |
|---|---------|---------|---------|---------|---------|---------|---------|
| | | Group 1 | Group 2 | Group 3 | Group 1 | Group 2 | Group 3 |
| % of budget spent externally on applied research | Inputs | | | | | O | |
| % of budget spent internally on applied research | Inputs | | | X | | | |
| % of budget spent internally on basic research | Inputs | | | X | | | |
| Amount of facilities like laboratories, libraries, and equipment employed | Inputs | | X | | O | | O |
| Annual spending for market investigations aimed at generating technological innovation | Inputs | | | | | O | |
| Degree of flexibility of resources employed | Inputs | | | | O | | |
| Employee retention rate | Inputs | | | | O | O | O |
| Hours spent on projects vs. total hours R&D | Inputs | | | X | | | |
| Innovation level and degree of creativity | Inputs | | X | X | O | | |
| Number of funds invested | Inputs | | | | | O | |
| Evaluation of scope of patent applications | R&D Lab | | | | | | O |
| Product development cycle time | R&D Lab | | | X | | | |
| Project progress and projects completed | R&D Lab | X | X | X | | | |
| Quality of project management | R&D Lab | | X | | | | |
| % of projects abandoned after a certain degree of completion | Outputs | X | X | | | | |
| % of projects in which customers are operatively involved (training, testing, or problem solving) | Outputs | | | | O | | |
| % of projects technically successful | Outputs | X | | | | | |
| Average cost of each completed project | Outputs | | | | O | | |
| Cost efficiency, i.e. keeping within budget | Outputs | | | X | | | |
| Degree of anticipation of internal customer needs | Outputs | X | X | | | | |
| Degree of project attractiveness | Outputs | | | | | O | |
| Design performance, i.e. as to manufacturing cost, manufacturability, or testability | Outputs | | | X | | O | |
| Number of new projects initiated | Outputs | X | | | | | |
| Number of parts produced | Outputs | | | | O | | O |
| Number of products designed | Outputs | X | | | | | |
| Number of redesigns and average time of redesigns | Outputs | | | | O | | O |
| Number of other legal protection methods besides patents | Outputs | | | | O | O | O |
| Percent of new technology content in new products | Outputs | X | | X | | O | |
| Planning accuracy, i.e. % of agreed milestones and/or objectives met | Outputs | | | X | | | |
| Project feasibility with milestones/gates | Outputs | | | | | | O |
| Quality, i.e. number of times reworked | Outputs | | X | | | | O |
| Rate of re-use of standard designs, i.e. as proven technology | Outputs | | X | | O | | O |
| Sum of planned or revised project durations | Outputs | | X | | | | O |
| Technology leadership in terms of number of patents developed | Outputs | | | | | O | |
| The ratio of cost to time and to time per technical progress unit | Outputs | | | | | O | |
| The ratio of technical progress to costs | Outputs | X | | | | | |
| The ratio of technical progress to time | Outputs | X | | | | | |
| Transfer rate of new knowledge and technology into product development | Outputs | X | X | | | | |

TABLE 5 Summary of key and flop selected measures

| Measure | Level | KEY_Score | Flop_Score |
|--|---------|-----------|------------|
| <i>Top selected</i> | | | |
| Project progress and projects completed | R&D Lab | 3 | 0 |
| % of projects abandoned after a certain degree of completion | Outputs | 2 | 0 |
| Degree of anticipation of internal customer needs | Outputs | 2 | 0 |
| Transfer rate of new knowledge and technology into product development | Outputs | 2 | 0 |
| Percent of new technology content in new products | Outputs | 2 | 1 |
| Innovation level and degree of creativity | Input | 2 | 1 |
| <i>Flop selected</i> | | | |
| Employee retention rate | Inputs | 0 | 3 |
| Number of other legal protection methods besides patents | Outputs | 0 | 3 |
| Number of parts produced | Outputs | 0 | 2 |
| Number of redesigns and average time of redesigns | Outputs | 0 | 2 |
| Rate of re-use of standard designs, i.e. as proven technology | Outputs | 1 | 2 |
| Amount of facilities like laboratories, libraries, and equipment employed | Input | 1 | 2 |
| <i>Thereof equivocal</i> | | | |
| Amount of facilities like laboratories, libraries, and equipment employed | Inputs | 1 | 2 |
| Innovation level and degree of creativity | Inputs | 2 | 1 |
| Design performance, i.e. as to manufacturing cost, manufacturability, or testability | Outputs | 1 | 1 |
| Percent of new technology content in new products | Outputs | 2 | 1 |
| Quality, i.e. number of times reworked | Outputs | 1 | 1 |
| Rate of re-use of standard designs, i.e. as proven technology | Outputs | 1 | 2 |
| Sum of planned or revised project durations | Outputs | 1 | 1 |

important. On the flop-ranked side, this is mirrored by *the ratio of cost to time and to time per technical progress unit* (month, year, etc.) (useful in the survey assessment). Only the reassessment through focus group 2 identified a related measure (*hours spent on projects vs. total*

TABLE 6 Summary of key and flop selected measures

| Top selected measures | Level | Focus group |
|--|---------|-------------|
| % of budget spent internally on applied research | Inputs | FG2 |
| % of budget spent internally on basic research | Inputs | FG2 |
| Hours spent on projects vs. total hours R&D | Inputs | FG2 |
| Innovation level and degree of creativity | Inputs | FG1 |
| Transfer rate of new knowledge and technology into product development | Outputs | FG1 + FG2 |
| % of projects abandoned after a certain degree of completion | Outputs | FG1 |
| Degree of anticipation of internal customer needs | Outputs | FG1 |
| Percent of new technology content in new products | Outputs | FG1 |
| Planning accuracy, i.e. % of agreed milestones and/or objectives met | Outputs | FG2 |
| Project progress and projects completed | R&D Lab | FG1 + FG2 |

hours R&D) as a key measure. This might give an indication of the frequency and timing of measure assessment: When should measures be assessed? Does an annual assessment trump a monthly assessment or even a dynamic one (Salimi & Rezaei, 2018)?

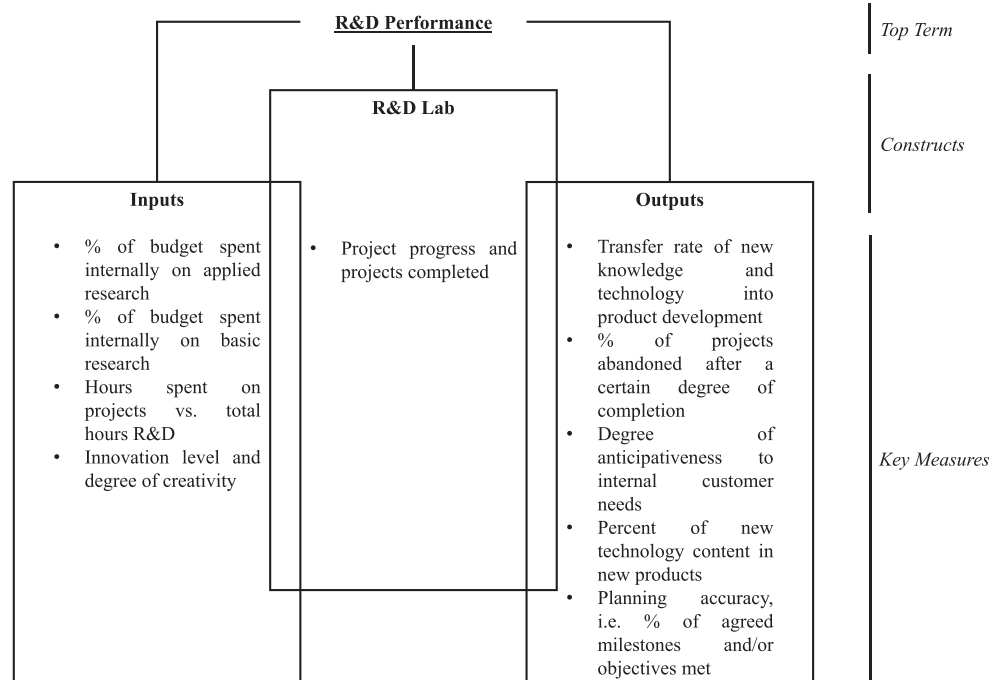
As with ideas, measures related to indicators like staff performance, data quality, or lead time (Wang et al., 2010), were not identified in our analyses. Ideas might be subsumed in the measure *innovation level and degree of creativity*, but a clear focus on measurements covering predictability and on-time schedule, as also proposed by Wang et al. (2010), is visible in the following key-selected measures: *project progress and projects completed* and *percentage of projects abandoned after a certain degree of completion*. It seems as if (1) quantity and (2) costs are perceived as more important than (3) quality, in some form ranking these three dimensions (e.g. by Brown & Svenson, 1998; Coccia, 2001) of measure evaluation. This is also reflected within the measures that were equivocally rated by the first focus group as both key and flop, like *innovation level and degree of creativity*, *design performance*, *percent of new technology content in new products*, and *quality, i.e. number of times reworked*. This confirms the findings from the analysis of typical output measures as proposed by Brown and Svenson (1998) with result-driven measures, e.g. project completion, trumping softer measures like scientific output and quality of scientific output.

In contrast to Brown and Svenson (1998), some of the measures like the *degree of anticipation of internal customer needs* were perceived both by the descriptive survey respondents and the focus groups as a key measure, indicating that behavioral measures should not be excluded per se. As with behavior, the characteristic of "focus on external vs. internal measurement" (Brown & Svenson, 1998) have been addressed in prior literature, especially when it comes to innovation network effects. Researchers stressed the importance of external validity of R&D performance or indicators to measure performance in relation to activities outside the R&D department. The results of our

TABLE 7 Reconciliation of R&D Lab measures to key performance measures

| Original R&D Lab measures (core measures bold) | Level | Focus group attribution |
|---|----------------|---|
| Researching | R&D Lab | % of budget spent internally on applied research |
| Developing | R&D Lab | % of budget spent internally on basic research |
| Funds invested | Input | |
| People/Number of hours worked | Input | Hours spent on projects vs. total hours R&D |
| Ideas | Input | Innovation level and degree of creativity |
| Results | R&D Lab | % of projects abandoned after a certain degree of completion |
| Milestones met | Outputs | Planning accuracy, i.e. % of agreed milestones and/or objectives met |
| Specific requests | Input | Degree of anticipation of internal customer needs |
| <i>Project progress/Projects completed</i> | <i>Outputs</i> | <i>Project progress and projects completed</i> |
| <i>Transfer rate of new knowledge and technology into product development</i> | <i>Outputs</i> | <i>Transfer rate of new knowledge and technology into product development</i> |
| <i>Percent of new technology content in new products</i> | <i>Outputs</i> | <i>Percent of new technology content in new products</i> |

FIGURE 4 R&D laboratory system extension



descriptive survey show that respondents did not deem any of the measures that have been identified in prior literature (as in Kerssens-van Drongelen & Bilderbeek, 1999; Chiesa & Masella, 1996; Wang et al., 2010; Lazzarotti et al., 2011) as useful or did not find them to be applied in their firms. Only three of these ten measures were known, namely *percentage of collaborative projects with third parties* (Wang et al., 2010), *number of alliances dedicated to technological innovation* (Lazzarotti et al., 2011), and *externally built network* (Kerssens-van Drongelen & Bilderbeek, 1999). This is surprising given the fact that most respondents' firms are mid-sized firms with an international business perspective engaging in various inbound and outbound R&D activities (as defined by Gassmann & Enkel, 2004). A possible explanation might be that people active in the R&D department are overly

focused on measures within their immediate work environment or might not be aware of the importance of measuring these external effects for department-centric performance. However, all industry experts participating in the expert survey and focus groups were members of an innovation network, hence were actively engaging in the exchange and profiting from network effects. Given this special innovation network relation, the lack of selection of network effect measures is even more surprising. Possibly, members of an innovation network perceive the significance of exchange and measurement at network level differently from exchange in actual R&D projects at the individual firm level. Nevertheless, the key identified measure *degree of anticipation of internal customer needs* might indicate that not all external measurements have a lower importance, but that at least

internal firm needs, i.e. external to the R&D department, might play a major role in assessing R&D performance or collaboration efficiency (Lee et al., 2019).

The findings of our study also contribute to the discussion on characteristics of successful measurement systems (Brown & Svenson, 1998; Chiesa et al., 2007). Depending on the objectives of measurement, characteristics of the respective system employ different design choices that can also lead to organizational benefits like resource allocation optimization (Bourne, Franco-Santos, Micheli, & Pavlov, 2018; Chiesa et al., 2007). In addition, these can contribute to employees feeling more connected to firm goals and performance measurement necessities (Englund & Ludvigsen, 2015).

The final selection of key performance measures indicates that input and output measures outnumber R&D Lab measures, supporting that “measur[ing] only valuable accomplishments/outputs” seems to be of key importance (Brown & Svenson, 1998). This is also in line with the characteristic of “focus[ing] on measuring outcomes and outputs, not behavior”, as outputs play a dominant role amongst the performance measures. We cannot conclude on the outcome characteristics, as our research focus was specified on measuring R&D performance only at the R&D department level, as depicted in steps 1–3 of the R&D Lab (Figure 1). Contrary to Brown and Svenson (1998), we refrained from including outcomes as a measure, as outcomes can only be assessed retrospectively and cannot serve as (forward-looking) R&D performance indicators. This theoretical foundation guiding our approach was subject to all focus group workshops and has been reflected and confirmed in our findings.

7 | MANAGERIAL IMPLICATIONS

Managers can assess their current status and improve their existing R&D performance measures to better connect business and innovation activities for an enhanced performance of the individual firm and the innovation network. The approach outlined in this paper can serve as a framework for arriving at meaningful R&D performance measures (see Table 8).

In a first step, managers should assess the status-quo, answering questions like which measures are currently being used. Building on this, measures which are not used should be identified to form a combined longlist of potential measures, as demonstrated through our review of prior literature. Next, these measures should be triangulated and scrutinized according to parameters like practical applicability and measurement focus, eventually resulting in a list of key- and flop-ranked measures. As mentioned by one participant, “during this process it is less relevant if 10, 20, or 30 measures are identified [...] since it gives us the potential to further discuss their usefulness” (Participant of focus group workshop 2, July 19, 2018). The identified measures could then be implemented and assessed over a predefined period of time under constant monitoring and refinement. By doing so, other uses of these measures than performance measurement alone, like suitability of implemented measures as controlling instruments, could improve the effectiveness and efficiency of the R&D

TABLE 8 Managerial framework for R&D performance measures

| For practice | Steps in this paper |
|--------------|---|
| Step 1 | Assess the status-quo Literature review Appendix |
| Step 2 | Identify unused measures |
| Step 3 | Form combined longlist of potential measures |
| Step 4 | Triangulated and scrutinized measures Textual analysis Table 2 |
| Step 5 | Identify key- and flop-ranked measures Survey, focus group 1 Tables 4 and 5 |
| Step 6 | Use as potential for further discussion Focus group 2 |
| Step 7 | Derivation of key performance measure Focus group 3 Table 6 |
| Step 8 | Implemented measures |
| Step 9 | Assess measures/constant monitoring |
| Step 10 | Measure refinement |
| Step 11 | Measure extension |

performance measures' evaluation process. This paper provides managers with a comprehensive list of measures (at various selection stages) that can be used as a starting point in evaluating innovation measures internally or in innovation network settings.

Our findings further indicate that managers should be aware of department-specific idiosyncrasies: Being overly focused on R&D department-centric performance measures that relate to the direct work environment, might cause less obvious measures to be overlooked, like measuring external or network effects, or measures related to human resource indicators. Resource constraints might further enhance the problem of evaluator-fit, i.e. which employees are best suited to measure R&D performance and draw inferences on measurement results.

Moreover, R&D performance measures might not only impact the individual firm. Our research approach and findings show how innovation networks can foster the identification and development of innovation performance measures (Agostini, Filippini, & Nosella, 2015). Network members enhanced the network performance by contributing proprietary, firm-specific knowledge to the network. These accumulated knowledge inputs contributed to creating proprietary network performance measures within the innovation network itself, as visible throughout our focus groups. The knowledge of these measures then transfers back to the network member firms, influencing and reinforcing performance measure initiatives. With this kind of innovation network, knowledge exchange flows in both directions, eventually creating performance measures on the individual as well as at the innovation network level.

In the case of the ratio of input and output measures among the key selected measures, an active exchange and information provision, as fostered through innovation networks, enhances the understanding of measures. This can lead to differing, more nuanced results (focus

group 1 selected almost no inputs measures, whereas, after inclusion of the results of focus group 2, 40 percent of key performance measures could be attributed to the input stage).

8 | LIMITATIONS AND FURTHER RESEARCH

Our methodological approach aimed to identify key performance measures. We primarily focused on a structured derivation of R&D performance measures from prior literature. Through a mixed-methods approach spanning textual analysis, triangulation, descriptive survey, and three focus group assessments, we summarized relevant key R&D performance measures. Even though our sample and basis of analysis consisted of more than 40 industry experts in the focus groups alone, this basis needs to be broadened with more data, potentially enabling further quantitative validation.

Even though the firms who participated were from a large variety of different industries and of sizes, additional (specific) industries could be included to further enhance the generalizability of our results. Another limitation is the time aspect. Although the information was collected at various points in time, a longitudinal study would allow for differing insights. In addition, all participating firms are located in Germany. This might limit the generalizability of our findings since the attitude and approach toward measures might be influenced by the culture of German firms. Future researchers should include a broader international background, thereby focusing on different cultural backgrounds.

Despite a potential indication of the focus groups that the time of assessment might not play an important role in R&D performance measurement, participants of the focus groups noticed that some measures might be more suitable at certain points in time (as also indicated by, e.g., Chiesa et al., 2007). Measures at the input stage are obviously more suitable to measure the R&D performance at the beginning of an R&D project. It is proposed that a dynamic measurement might be most suited, mirroring prior research that recommends some form of index to measure R&D performance (Birchall et al., 2011; Brown & Svenson, 1998). Future research could assess our findings as to whether they form a suitable basis for a dynamic measurement process throughout the full R&D project cycle, i.e. spanning all stages from inputs, over R&D Lab and outputs, while dynamically adjusting to the specific requirements at and throughout each stage. The proposed hybrid nature of some input measures, i.e. qualifying both as input and output measures alike, could be of importance in these considerations. Potential effects of these measures should be included in elaborations on dynamic measures in further research.

This reflects on questions of practicability of the identified measures. As one innovation manager noticed: "This is the tricky part: I always wonder how someone, as innovation manager in R&D, can access these controlling-based figures?" (Participant of focus group workshop 1, December 8, 2017). Reinforcing this notion, our findings revealed a trend towards outputs and project-related measures. Future research should analyze these trends in the light of applicable

perspectives of performance, as proposed by Wang et al. (2010) and Lazzarotti et al. (2011). This might shed light on how to apply these measures in practice, how the R&D department might include and access these measures, and which evaluators are best suited to collect, analyze, and decide on the measures' effects.

ACKNOWLEDGMENTS

The authors would like to thank all participants who contributed in the various workshops for this article. Their engagement and openness is highly appreciated.

ORCID

Peter M. Bican  <https://orcid.org/0000-0003-4937-5818>

Alexander Brem  <https://orcid.org/0000-0002-6901-7498>

REFERENCES

- Agostini, L., Filippini, R., & Nosella, A. (2015). Management and performance of strategic multipartner SME networks. *International Journal of Production Economics*, 169, 376–390.
- Azzone, G., & Maccarrone, P. (1997). The emerging role of lean infrastructures in technology transfer: The case of the Innovation Plaza project. *Technovation*, 17, 391–402.
- Baglieri, E., Chiesa, V., Grando, A., & Manzini, R. (2001). Evaluating intangible assets: the measurement of R&D performance. Research Division Working Paper No. 01/49. Available at: <https://doi.org/10.2139/ssrn.278260>
- Beers, C., & Zand, F. (2014). R&D cooperation, partner diversity, and innovation performance. *Journal of Product Innovation Management*, 31, 292–312.
- Bican, P. M., Guderian, C. C., & Ringbeck, A. (2017). Managing knowledge in open innovation processes: An intellectual property perspective. *Journal of Knowledge Management*, 21, 1384–1405.
- Birchall, D., Chanaron, J.-J., Tovstiga, G., & Hillenbrand, C. (2011). Innovation performance measurement: Current practices, issues and management challenges. *International Journal of Technology Management*, 56, 1–20.
- Black, T. R. (1993). *Evaluating social science research: An introduction*. London: Sage.
- Blindenbach-Driessen, F., van Dalen, J., & van den Ende, J. (2010). Subjective performance assessment of innovation projects. *Journal of Product Innovation Management*, 27, 572–592.
- Bourne, M., Franco-Santos, M., Micheli, P., & Pavlov, A. (2018). Performance measurement and management: A system of systems perspective. *International Journal of Production Research*, 56, 2788–2799.
- Branley, D., Seale, C., & Zacharias, T. (2018). Doing a literature review. 2In C. Seale (Ed.), *Researching society and culture* (4th ed.). London: Sage.
- Bremser, W. G., & Barsky, N. P. (2004). Utilizing the balanced scorecard for R&D performance measurement. *R&D Management*, 34, 229–238.
- Brown, M. G., & Svenson, R. A. (1998). Measuring R&D productivity. *Research-Technology Management*, 41(6), 30–35.
- Chesbrough, H. (2003). *Open innovation: The new imperative for creating and profiting from technology*. Cambridge, MA: Harvard Business School Press.
- 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, 283–301.
- Chiesa, V., Frattini, F., Lazzarotti, V., & 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, 187–202.

- 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, 487–519.
- Chiesa, V., & Masella, C. (1996). Searching for an effective measure of R&D performance. *Management Decision*, 34(7), 49–57.
- Coccia, M. (2001). A basic model for evaluating R&D performance: Theory and application in Italy. *R&D Management*, 31, 453–464.
- Conley, J. G., Bican, P. M., & Ernst, H. (2013). Value articulation: A framework for the strategic management of intellectual property. *California Management Review*, 55(4), 102–120.
- Cooper, R. G., & Kleinschmidt, E. J. (1995). Benchmarking the firm's critical success factors in new product development. *Journal of Product Innovation Management*, 12, 374–391.
- Denzin, N. K. (1978). *Sociological methods: A sourcebook* (2nd ed.). New York: McGraw-Hill.
- Detzen, N., Verbeeten, F. H. M., Gamm, N., & Möller, K. (2018). Formal controls and team adaptability in new product development projects. *Management Decision*, 56, 1541–1558.
- Dziallas, M., & Blind, K. (2019). Innovation indicators throughout the innovation process: An extensive literature analysis. *Technovation*, 80–81, 3–29.
- Englund, H., & Ludvigsen, E. (2015). *Performance measurement systems as management control in R&D organizations: A case study*. Master's Thesis. Karlskrona, Sweden: Blekinge Tekniska Högskola.
- Enkel, E., Gassmann, O., & Chesbrough, H. (2009). Open R&D and open innovation: Exploring the phenomenon. *R&D Management*, 39, 311–316.
- Freeman, T. (2006). "Best practice" in focus group research: Making sense of different views. *Journal of Advanced Nursing*, 56, 491–497.
- Gassmann, O., & Enkel, E. (2004). Towards a theory of open innovation: Three core process archetypes. *Proceedings of the R&D Management Conference*, Lisbon, July 6–9, 2004.
- Gray, D. E. (2018). *Doing research in the real world* (4th ed.). London: Sage.
- Griffin, A. (1997). PDMA research on new product development practices: Updating trends and benchmarking best practices. *Journal of Product Innovation Management*, 14, 429–458.
- Hanson, W. E., Creswell, J. W., Clark, V. L. P., Petska, K. S., & Creswell, J. D. (2005). Mixed methods research designs in counseling psychology. *Journal of Counseling Psychology*, 52, 224–235.
- Hart, C. (2018). *Doing a literature review: Releasing the research imagination* (2nd ed.). London: Sage.
- Henttonen, K., Ojanen, V., & Puumalainen, K. (2016). Searching for appropriate performance measures for innovation and development projects. *R&D Management*, 46, 914–927.
- Hienrath, C. and Riar, F. (2015). The architecture of evaluation processes in open innovation settings. Paper presented at the 75th Academy of Management conference, Vancouver, Canada.
- Jesson, J., Matheson, L., & Lacey, F. M. (2011). *Doing your literature review: Traditional and systematic techniques*. London: Sage.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14–26.
- Kale, P., & Singh, H. (2009). Managing strategic alliances: What do we know now, and where do we go from here? *Academy of Management Perspectives*, 23(3), 45–62.
- Kerssens-van Drongelen, & Bilderbeek, J. (1999). R&D performance measurement: More than choosing a set of metrics. *R&D Management*, 29, 35–46.
- Kerssens-van Drongelen, & Cooke, A. (1997). Design principles for the development of measurement systems for research and development processes. *R&D Management*, 27, 345–357.
- Khanna, R., Guler, I., & Nerkar, A. (2016). Fail often, fail big, and fail fast? Learning from small failures and R&D performance in the pharmaceutical industry. *Academy of Management Journal*, 59, 436–459.
- Kristiansen, J. N., & Ritala, P. (2018). Measuring radical innovation project success: Typical metrics don't work. *Journal of Business Strategy*, 39(4), 34–41.
- Lakiza, V., & Deschamps, I. (2018). How to develop innovation KPIs in an execution-oriented company. *Technology Innovation Management Review*, 8(7), 14–30.
- Larsen, A., & Lindquist, P. (2016). *A performance measurement framework for R&D activities: Increasing transparency of R&D value contribution*. MSc thesis. Stockholm, Sweden: KTH Industrial Engineering and Management.
- Lazzarotti, V., Manzini, R., & Mari, L. (2011). A model for R&D performance measurement. *International Journal of Production Economics*, 134, 212–223.
- Lee, J., Kim, C., & Choi, G. (2019). Exploring data envelopment analysis for measuring collaborated innovation efficiency of small and medium-sized enterprises in Korea. *European Journal of Operational Research*, 278, 533–545.
- McKim, C. A. (2017). The value of mixed methods research: A mixed methods study. *Journal of Mixed Methods Research*, 11, 202–222.
- Menezes, L., Sellitto, M., Librelato, T., Borchardt, M., & Pereira, G. (2016). Identification and quantification of influent factors in perceived quality of the e-service provided by a university. *Business Process Management Journal*, 22, 438–457.
- Neely, A., Mills, J., Platts, K., Gregory, M., & Richards, H. (1996). Performance measurement system design: Should process based approaches be adopted? *International Journal of Production Economics*, 46–47, 423–431.
- Perkmann, M., Neely, A., & Walsh, K. (2011). How should firms evaluate success in university–industry alliances? *R&D Management*, 41, 202–216.
- Salimi, N., & Rezaei, J. (2018). Evaluating firms' R&D performance using best worst method. *Evaluation and Program Planning*, 66, 147–155.
- Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 69, 107–131.
- Seal, D. W., Bogart, L. M., & Ehrhardt, A. A. (1998). Small group dynamics: The utility of focus group discussions as a research method. *Group Dynamics: Theory, Research, and Practice*, 2, 253–266.
- Sellitto, M. A. (2018). Assessment of the effectiveness of green practices in the management of two supply chains. *Business Process Management Journal*, 24, 23–48.
- Sellitto, M. A., & Luchese, J. (2018). Systemic cooperative actions among competitors: The case of a furniture cluster in Brazil. *Journal of Industry, Competition and Trade*, 18, 513–528.
- Serrano-Bedia, A. M., López-Fernández, M. C., & García-Piqueres, G. (2018). Complementarity between innovation knowledge sources: Does the innovation performance measure matter? *Business Research Quarterly*, 21, 53–67.
- Škec, S., Cash, P., & Štorga, M. (2017). A dynamic approach to real-time performance measurement in design projects. *Journal of Engineering Design*, 28, 255–286.
- Stewart, D. W., & Shamdasani, P. N. (2015). *Focus groups: Theory and practice* (3rd ed.). London: Sage.
- Teece, D. J. (2007). Explicating dynamic capabilities: The nature and micro-foundations of (sustainable) enterprise performance. *Strategic Management Journal*, 28, 1319–1350.
- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18, 509–533.
- Tkocz, A., Munck, J. C., & Wald, A. E. (2018). Innovation management control: Bibliometric analysis of its emergence and evolution as a research field. *International Journal of Innovation Management*, 22(3), 1850031.
- Wang, J., Lin, W., & Huang, Y.-H. (2010). A performance-oriented risk management framework for innovative R&D projects. *Technovation*, 30, 601–611.

- Werner, B. M., & Souder, W. E. (1997). Measuring R&D performance—State of the art. *Research-Technology Management*, 40(2), 34–42.
- Wheelwright, S. C., & Clark, K. B. (1992). Creating project plans to focus product development. *Harvard Business Review*, 70(2), 70–82.

AUTHOR BIOGRAPHIES

Peter M. Bican is Assistant Professor at the Chair of Technology Management at Friedrich-Alexander-University Erlangen-Nuremberg (Germany). His research focus is on strategic asset management in innovation and technology management.

Alexander Brem is Endowed Chaired Professor and Institute Head at the University of Stuttgart (Germany). In addition, he is Honorary Professor at the University of Southern Denmark (Denmark). His research focus is on technological innovation and entrepreneurship.

How to cite this article: Bican PM, Brem A. Managing innovation performance: Results from an industry-spanning explorative study on R&D key measures. *Creat Innov Manag.* 2020;29:268–291. <https://doi.org/10.1111/caim.12370>

APPENDIX

Measures identified from literature

| Measure | Source | Measure | Source |
|---|--|--|--|
| % of budget spent externally on applied research | Kerssens-van Drongelen & Bilderbeek (1999) | Designs produced | Brown & Svenson (1998) |
| % of budget spent externally on basic research | Kerssens-van Drongelen & Bilderbeek (1999) | Developing | Brown & Svenson (1998) |
| % of budget spent internally on applied research | Kerssens-van Drongelen & Bilderbeek (1999); Chiesa & Frattini (2007) | Discounted cash flows | Baglieri et al. (2001) |
| % of budget spent internally on basic research | Kerssens-van Drongelen & Bilderbeek (1999); Chiesa & Frattini (2007) | (Drug-related) adverse events and efficacy | Wang et al. (2010) |
| % of evaluation ideas applied in new projects | Wang et al. (2010); Chiesa et al. (2009) | Efficiency/keeping within budget | Kerssens-van Drongelen & Bilderbeek (1999); Chiesa et al. (2007) |
| % of products succeeding in the market | Kerssens-van Drongelen & Bilderbeek (1999); Chiesa & Frattini (2007) | Employee retention rate | Bremser & Barsky (2004); Wang et al. (2010) |
| % of project evaluation ideas applied in new projects | Kerssens-van Drongelen & Bilderbeek (1999) | Equipment | Brown & Svenson (1998); Coccia (2001) |
| % of projects abandoned after a certain degree of completion | Chiesa & Masella (1996) | Expected or realized IRR/ROI | Kerssens-van Drongelen & Bilderbeek (1999); Chiesa et al. (2009) |
| % of projects technically successful | Chiesa & Masella (1996) | Expected royalties | Baglieri et al. (2001) |
| % of sales and profits from new products introduced in the last three (five) years | Chiesa & Masella (1996); Werner & Souder (1997) | Facilities/Laboratories | Brown & Svenson (1998); Coccia (2001) |
| % of sales and profits from products with significant enhancements in the last three (five) years | Chiesa & Masella (1996) | Facts | Brown & Svenson (1998) |
| % of sales by new product(s) | Kerssens-van Drongelen & Bilderbeek (1999) | Flexibility of resources | Baglieri et al. (2001) |
| Accuracy of pricing and revenue planning | Wang et al. (2010) | Funds invested | Brown & Svenson (1998) |
| Agreed milestones/objectives met | Kerssens-van Drongelen & Bilderbeek (1999); Baglieri et al. (2001) | High design for manufacturability | Kerssens-van Drongelen & Bilderbeek (1999) |
| Alignment to business strategy | Baglieri et al. (2001) | Hours spent on projects/total hours R&D | Kerssens-van Drongelen & Bilderbeek (1999); Chiesa et al. (2009) |
| Annual spending for market investigations aimed at generating technological innovation | Lazarrotti et al. (2011) | Human capital indicators | Perkmann et al. (2011) |
| Anticipativeness to external customer needs | Werner & Souder (1997); Kerssens-van Drongelen & Bilderbeek (1999) | Ideas | Brown & Svenson (1998); Kerssens-van Drongelen & Bilderbeek (1999); Chiesa et al. (2009) |
| Anticipativeness to internal customer needs | Kerssens-van Drongelen & Bilderbeek (1999) | In-process quality control tool | Brown & Svenson (1998); Larsen & Lindquist (2016) |
| Approval ratio of new (drug) | Wang et al. (2010) | Incremental profitability | Baglieri et al. (2001) |

| Measure | Source | Measure | Source |
|---|---|---|--|
| Average cost of each completed project | Lazzarotti et al. (2011) | Industry income | Perkmann et al. (2011) |
| Average development cost per new products | Wang et al. (2010) | Information | Brown & Svenson (1998); Brown & Svenson (1998) |
| Average development cycle time | Wang et al. (2010) | Interaction intensity/Relevant research | Perkmann et al. (2011) |
| Awards won | Brown & Svenson (1998) | Joint objective setting/High quality research | Perkmann et al. (2011) |
| Behavior, i.e. a process or activity/professional esteem | Brown & Svenson (1998); Kerssens-van Drongelen & Bilderbeek (1999) | Knowledge | Brown & Svenson (1998); Kerssens-van Drongelen & Bilderbeek (1999); Chiesa et al. (2009) |
| Capacity for self-financing | Coccia (2001) | Learning organization | Kerssens-van Drongelen & Bilderbeek (1999) |
| Combine data with external developments | Kerssens-van Drongelen & Bilderbeek (1999) | Leverage factor/Funding | Perkmann et al. (2011) |
| Combine data with internal conditions | Kerssens-van Drongelen & Bilderbeek (1999) | Libraries | Coccia (2001) |
| Committee work | Brown & Svenson (1998) | Life cycle cost | Baglieri et al. (2001); Chiesa et al. (2007) |
| Competitive value of technology innovations and research data | Brown & Svenson (1998) | Long-term focus | Kerssens-van Drongelen & Bilderbeek (1999) |
| Cost of loss | Baglieri et al. (2001) | Long-term visibility | Baglieri et al. (2001) |
| Creativity/innovation level | Kerssens-van Drongelen & Bilderbeek (1999); Lakiza & Deschamps (2018) | Market share gained due to R&D | Kerssens-van Drongelen & Bilderbeek (1999) |
| Credit | Coccia (2001) | Market share of new product | Wang et al. (2010) |
| Current percentage of sales of new products | Wang et al. (2010) | Milestones met | Brown & Svenson (1998); Chiesa & Masella (1996) |
| Current time to market/reference time to market | Kerssens-van Drongelen & Bilderbeek (1999) | National insurance coverage | Wang et al. (2010) |
| Customer lifetime value | Wang et al. (2010); Lakiza & Deschamps (2018) | Net income from sales of technology | Brown & Svenson (1998) |
| Customer satisfaction (functional product performance, product range, product variety, quality, etc.) | Chiesa & Masella (1996) | Net present cash flow to development cost | Brown & Svenson (1998) |
| Customer satisfaction with new products | Wang et al. (2010) | Net present value | Baglieri et al. (2001) |
| Deadlines | Brown & Svenson (1998); Chiesa & Masella (1996) | Network building | Kerssens-van Drongelen & Bilderbeek (1999) |
| Dependence on centralized funding | Coccia (2001) | | |
| New projects | Perkmann et al. (2011) | R&D productivity, which influences cash outflows occurring before t^* | Chiesa & Masella (1996) |
| No. of innovations—Impact on ongoing R&D programs | Perkmann et al. (2011) | R&D projects which led to new or enhanced product or process innovations, licenses, patents | Chiesa & Masella (1996) |
| No. of solution concepts | Perkmann et al. (2011) | R&D revenue/R&D expenditure | Wang et al. (2010) |
| No. of times reworked | Kerssens-van Drongelen & Bilderbeek (1999) | Rate of re-use of standard designs/proven technology | Kerssens-van Drongelen & Bilderbeek (1999) |
| Number of alliances dedicated to technological innovation | Lazzarotti et al. (2011) | Real options | Baglieri et al. (2001) |

(Continues)

| | | | |
|--|--|---|--|
| Number of employees dedicated to external relationships in R&D | Lazzarotti et al. (2011); Chiesa et al. (2009) | Reproduction cost | Baglieri et al. (2001) |
| Number of parts produced | Brown & Svenson (1998) | Researching | Brown & Svenson (1998) |
| Number of personnel in training | Coccia (2001); Wang et al. (2010) | Results | Brown & Svenson (1998) |
| Number of redesigns and average time of redesigns; design performance (manufacturing cost, manufacturability, testability) | Chiesa & Masella (1996) | Sales derived from innovation projects | Lazzarotti et al. (2011); Lakiza & Deschamps (2018) |
| Number of reports/Documentation | Brown & Svenson (1998); Baglieri et al. (2001); Lazzarotti et al. (2011) | Sales from new products | Bremser & Barsky (2004); Wang et al. (2010); Larsen & Lindquist (2016) |
| Number of teaching courses held | Coccia (2001) | Scope of applications | Baglieri et al. (2001) |
| Number of technology transfer plans | Brown & Svenson (1998) | Short-term capability | Baglieri et al. (2001) |
| Outside data, i.e. other department than R&D | Brown & Svenson (1998) | Specific requests | Brown & Svenson (1998) |
| Patents/Legal protection | Brown & Svenson (1998); Coccia (2001); Kerssens-van Drongelen & Bilderbeek (1999); Serrano-Bedia et al. (2018) | Speed | Kerssens-van Drongelen & Bilderbeek (1999) |
| People/Number of hours worked | Brown & Svenson (1998); Coccia (2001); Wang et al. (2010) | Speed to market | Kerssens-van Drongelen & Bilderbeek (1999) |
| Percent of net income from products developed in the last five years | Brown & Svenson (1998) | Staff learning metrics | Bremser & Barsky (2004); Perkmann et al. (2011) |
| Percent of new technology content in new products | Brown & Svenson (1998) | Strategic skills coverage ratio | Wang et al. (2010) |
| Percent of products developed in the last five years | Brown & Svenson (1998) | Sum of revised project durations/Sum of planned durations | Kerssens-van Drongelen & Bilderbeek (1999); Chiesa et al. (2009) |
| Percentage of collaborative projects with third parties | Wang et al. (2010); Kerssens-van Drongelen & Bilderbeek (1999) | Technological excellence | Baglieri et al. (2001); Baglieri et al. (2001) |
| Percentage of people having a pertinent degree (with respect to all the people devoted to technological innovation) | Lazzarotti et al. (2011) | Technology leadership | Kerssens-van Drongelen & Bilderbeek (1999) |
| Percentage of projects in which customers are operatively involved (training, test, problem solving) | Lazzarotti et al. (2011) | Technology/design re-use | Kerssens-van Drongelen & Bilderbeek (1999) |
| Personalized product treatment | Wang et al. (2010) | Testing | Brown & Svenson (1998) |
| Planning accuracy | Kerssens-van Drongelen & Bilderbeek (1999) | The degree of integration with the manufacturing function, which influences costs after t^* | Chiesa & Masella (1996) |
| Presentations made | Brown & Svenson (1998) | The degree of integration with the marketing function | Chiesa & Masella (1996) |
| Principles | Brown & Svenson (1998) | The number of duplications | Chiesa & Masella (1996) |
| Processes | Brown & Svenson (1998); Serrano-Bedia et al. (2018) | The number of project parts jointly carried out with other projects | Chiesa & Masella (1996) |
| Product development cycle time | Brown & Svenson (1998); Chiesa & Frattini (2007) | The ratio of cost to technical progress that measures the cost per unit of technical progress | Chiesa & Masella (1996) |
| Product life cycle | Brown & Svenson (1998); Chiesa et al. (2009) | The ratio of cost to time and to time per technical progress unit | Chiesa & Masella (1996) |

| | | | |
|--|--|--|--|
| Products | Brown & Svenson (1998); Kerssens-van Drongelen & Bilderbeek (1999); Serrano-Bedia et al. (2018) | The ratio of technical progress to costs | Chiesa & Masella (1996); Werner & Souder (1997) |
| Products designed | Brown & Svenson (1998) | The ratio of technical progress to time, i.e. the technical progress per time unit (month, year, etc.) | Chiesa & Masella (1996) |
| Profit due to R&D | Kerssens-van Drongelen & Bilderbeek (1999) | The ratio of technical progress to time | Chiesa & Masella (1996) |
| Project attractiveness | Kerssens-van Drongelen & Bilderbeek (1999) | The ratio of time to technical progress, i.e. the time employed to improve technical progress unit | Chiesa & Masella (1996) |
| Project feasibility with milestones/gates | Kerssens-van Drongelen & Bilderbeek (1999) | Time dedicated to the analysis of reasons for failure of previous projects | Lazzarotti et al. (2011) |
| Project progress/Projects completed | Brown & Svenson (1998); Kerssens-van Drongelen & Bilderbeek (1999); Chiesa & Masella (1996) | Time to market | Bremser & Barsky (2004); Lazzarotti et al. (2011); Lakiza & Deschamps (2018) |
| Publications/Number of publications (books and paper) | Brown & Svenson (1998); Coccia (2001); Baglieri et al. (2001) | Time to market—Production costs savings | Baglieri et al. (2001); Chiesa & Frattini (2007) |
| Quality of new (drug) | Chiesa et al. (2007); Wang et al. (2010) | Time to market: Average concept to launch time, time for each phase (concept, design, initial production, launch) | Chiesa & Masella (1996) |
| Quality of project management | Baglieri et al. (2001); Chiesa et al. (2007); Larsen & Lindquist (2016) | Transfer rate of new knowledge and technology into product development | Brown & Svenson (1998); Coccia (2001) |
| R&D annual spending | Lazzarotti et al. (2011) | Updated cost | Baglieri et al. (2001) |
| R&D hit rate | Kerssens-van Drongelen & Bilderbeek (1999) | | |