Research Center
Finance & Information Management

Fraunhofer FIT
Project Group
Business & Information
Systems Engineering

Assessing IT Management's Performance: A Design Theory for Strategic IT Benchmarking

by

Katharina Ebner[1], Benjamin Mueller[2], Nils Urbach, Gerold Riempp[3], Helmut Krcmar[4]

in: IEEE Transactions on Engineering Management, 63, 1, 2016, p. 113-126

[1] University of Hagen
[2] University of Groningen
[3] EBS University of Business and Law
[4] Technische Universität München (TUM)

WI-539

University of Augsburg, D-86135 Augsburg
Visitors: Universitätsstr. 12, 86159 Augsburg
Phone: +49 821 598-4801 (Fax: -4899)

University of Bayreuth, D-95440 Bayreuth
Visitors: Wittelsbacherring 10, 95444 Bayreuth
Phone: +49 921 55-4710 (Fax: -844710)

www.fim-rc.de

Assessing IT Management's Performance: A Design Theory for Strategic IT Benchmarking

Authors

Katharina Ebner
Faculty of Business Administration and Economics
University of Hagen
katharina.ebner@fernuni-hagen.de

Benjamin Mueller (corresponding author)
Faculty of Economics and Business
University of Groningen
b.mueller@rug.nl

Nils Urbach
Faculty of Law, Business and Economics
University of Bayreuth
nils.urbach@uni-bayreuth.de

Gerold Riempp
EBS Business School
EBS University of Business and Law
gerold.riempp@ebs.edu

Helmut Krcmar
Faculty of Informatics
Technische Universität München (TUM)
krcmar@in.tum.de

Biographies

Dr. Katharina Ebner is a postdoctoral researcher at the Chair of Business Information Systems, Faculty of Business Administration and Economics at the University of Hagen. She received her PhD from the EBS Business School in 2014 and holds a diploma in Computer Science from Friedrich Schiller University Jena. Katharina's research focuses on the use and success of information systems, the digitalization of society and organizations, and benchmarking. Katharina has published her research in various outlets, e.g. the European Conference on Information Systems or the Hawaii International Conference on System Sciences. She serves the IS community as reviewer and editorial board member for various conferences, e.g. International and European Conference on Information Systems, and is a reviewer for the Journal of the AIS.

Dr. Benjamin Müller is an Assistant Professor at the University of Groningen, the Netherlands, and an Associate Researcher at the Institute for Enterprise Systems at the University of Mannheim, Germany. He received his PhD from the EBS Business School in 2010 and holds graduate degrees in Management and Information Systems from EBS Business School and Georgia State University. Benjamin's research is particularly focused on investigating how individual, organizational, and technological aspects of Enterprise Systems adapt to each other in times of organizational transformation. Complementary to that, he contributes the ongoing debate on theory and theorizing in the IS discipline. Benjamin has published over 40 peer-reviewed papers in various outlets such as, for example, the Journal of Management Information Systems, the Journal of the Association for Information Systems, or the International Conference on Information Systems. Benjamin serves the IS community in various editorial capacities.

Dr. Nils Urbach is a Professor of Information Systems and Strategic IT Management at the University of Bayreuth. Furthermore, he is Deputy Director of the Finance & Information Management (FIM) Research Center and the Project Group Business & Information Systems Engineering of Fraunhofer FIT. Before, he was Assistant Professor at EBS Business School in Wiesbaden. He also received his doctorate from EBS. Furthermore, he holds a diploma in information systems from the University of Paderborn. He gathered international experience during his research stays at the University of Pittsburgh and at Université de Lausanne. Complementary to his academic work, he was Consultant with Horváth & Partners and Accenture. His work has been published in several international journals such as the Journal of Information Technology, Journal of Strategic Information Systems, and Information & Management as well as in the proceedings of key international conferences.

Dr. Gerold Riempp is a full professor of information systems at EBS University of Business and Law in Wiesbaden, Germany. Before that, he had appointments as professor at Strasbourg, France and as senior lecturer at St. Gallen, Switzerland. His main fields of interest are workflow management, knowledge management, strategic IT management, and customer relationship management. In more than 20 years of academic research, Gerold has published numerous peer-reviewed articles in international conference proceedings and journals, among them the Journal of Management Information Systems and the Journal of Strategic Information Systems. Besides his academic career, he worked as senior manager, principal and partner in several international consulting firms such as PricewaterhouseCoopers, IMG and Horváth & Partners. In recognition of his dual qualification as both scholar and manager, he has been appointed as member and president of the supervisory board of several international IT and consulting companies.

Dr. Helmut Krcmar is professor and chair for information systems at the Technical University of Munich, Germany. He is a member of the TUM School of Management and was Dean of the Department of Informatics from 2010 to 2013. He studied business management in Saarbrücken, Germany and obtained his doctorate in 1983. After that, he worked as a postdoctoral fellow at the IBM Los Angeles Scientific Center and as an assistant professor of information systems at the Leonard Stern Graduate School of Business, New York University and at Baruch College, City University of New York. From 1987 to 2002, he held the Chair of Information Systems at the University of Hohenheim, Germany. He served as the AIS' president from 2014 to 2015. His research covers information management, IT-enabled value networks, service management, computer-supported cooperative work, and information systems for IT service providers in healthcare and e-government.

Assessing IT Management's Performance: A Design Theory for Strategic IT Benchmarking

Abstract

Given the continued economic pressure on IT organizations, the effective and efficient delivery of IT remains a crucial issue for IT executives in order to optimize their department's performance. Due to company specifics, however, an absolute assessment of IT organizations' performance has often proven difficult in the past. Consequently, many IT executives revert to comparative assessments such as IT benchmarking. IT benchmarking has been increasingly used to support IT management, also on a strategic level. While past research on such strategic IT benchmarking (SITBM) often focused on process models and optimal peer group composition, many practitioners repeatedly report problems with identifying or developing suitable methods for collecting the data needed for SITBM. We introduce a design theory for SITBM methods and illustrate how we derived the theory from an SITBM project over a period of 10 years. During that time, the method from which we abstract our design theory was applied in 102 different companies. We contribute to practice and theory by not only reporting a field-tested method for SITBM, but also by providing a design theoretical basis on how to develop such a method as a blueprint for IT managers that want to set-up an SITBM.

Keywords

Strategic IT Benchmarking, Strategic IT Management, IT Performance Measurement, Design Theory

1. Introduction

Chief information officers (CIOs) are required to justify service delivery costs and show their value contribution. As a consequence, they continually need appropriate information to assess their IT organization. In light of company specifics, it is often difficult to arrive at an absolute assessment. Strategic IT benchmarking (SITBM) has established itself as a preferred tool for measuring IT efficiency [1, 2], evaluating business-IT alignment [3–5], and IT strategy [6, 7] because it allows a structured comparison between IT management (ITM) and a peer group at a level of detail suitable for IT executives [1, 8]. The usefulness and quality of the data used for SITBM are of utmost importance to draw reliable conclusions. Past research has identified various facets contributing to the quality and usefulness of SITBM data. Among these facets are top management support [6, 9], an adequate process model [10–12], a sound peer group [11, 13–15], and a suitable data collection instrument [11, 16]. Although ITM researchers have suggested analyzing an IT organization's external environment, they most often focus on how to analyze the data and not on how to obtain it [17–19]. However, a lack of information and knowledge on known capabilities has been identified as a core problem in IT practitioners' strategic planning activities [17, 20]. While prior research in this area has rarely documented SITBM instruments and their development [10], some researchers have reported problems with SITBM that can be ascribed to insufficient instruments. For instance, many IT managers measure irrelevant information because they revert to "what can be measured, rather than what should be measured" [7, p. 21], or they have difficulty determining the relevant content [3, 8, 18]. Thus, desired strategic insights are difficult to derive [20]. Further, the validity of the benchmarking results is often corrupted due to deficient operationalizations and imprecise or too technical definitions [21, 22].

A challenge in developing SITBM instruments is that established guidelines for scientific measurement instrument development [e.g. 23, 24] are not applicable in this context because the nature and purpose of scientific and benchmarking instruments are very different. On the one hand, scientific instruments aim at the measurement of latent variables, which result from theories and hypotheses that are tested against observations or experimentations [23, p. 262]. Accordingly, guidelines for developing scientific instruments aim at developing and validating item sets and corresponding scales to measure latent variables [24]. By contrast, on the other hand, the ultimate goals of benchmarking instruments are comparison and data aggregation to make the instruments as parsimonious as possible while keeping them comprehensive [11, 25]. Accordingly, three relevant design goals can be identified. First, identification and mathematical operationalization of the performance indicators (PI) considered relevant *by practitioners* is targeted. These indicators are not latent and can typically be measured directly with one item [8]. Second, aggregation and normalization rules to be applied during data collection have to be identified to allow comparisons of different companies [1, 16]. Third, the collected data needs to be framed so that it can be reflected back to the IT organization [6, 7]. Traditional item development guidelines are not sufficient to guide the development of SITBM instruments.

In this paper, we report on a 10-year research project we conducted as part of an SITBM initiative involving 102 companies. Our research allows us to gain insights into how SITBM instruments should be designed to meaningfully support ITM activities. Moreover, our observations suggest that a stand-alone instrument is insufficient for SITBM, but also a glossary, a process model, and a role model are needed for successful application; a collection we refer to as an SITBM method. This paper develops a design theory (DT) for SITBM methods, referring to the eight components of an information systems design theory as proposed by Gregor and Jones [26]. The class of artifacts that can be instantiated from our DT are "methods for IS evaluation"

[27, p. 337]. The aim of these methods is to help IT executives systematically and effectively search the various strategic options in order to revise IT strategies.

To introduce our DT, this paper is structured as follows. In Section 2, we present the theoretical background that informs our research, particularly concerning the design of SITBM instruments, and derive an initial set of propositions and design principles (DPs). In Section 3, we introduce our research method. In Section 4, we provide a detailed presentation of our DT's iterative development and evaluation process, which has gone through a total of five iterations. In these iterations, 102 companies' feedback provides data for the iterative refinement of instrument, glossary, process model, and role model which we then abstract into our DT. As presenting all four parts would overload the paper, we focus on the instrument part and arrange the three other parts around this focus. We opt for a construction-oriented approach since SITBM has a practical application realm whose ultimate goal should be the development of useful artifacts. We describe the complete DT (including the revised set of initial propositions and DPs from Section 2) and the resulting instrument in Sections 5 and 6. In Section 7, we discuss and conclude our research.

2. Foundations

2.1 Structuring SITBM instruments

SITBM is strategic benchmarking applied to ITM. Thus, ITM provides the relevant content to be captured in SITBM and the context in which the instrument will be applied [8, 18]. The data needed for SITBM are gathered by instruments measuring the qualitative and the quantitative indicators of the ITM domains for which improvements are targeted. Prior research has documented only a few instruments applicable for SITBM. For instance, Saunders and Jones [28] suggest covering the costs and structures of the most important IT products and services while other authors maintain that an IT organization's key processes and structures should be considered (e.g. architecture or sourcing processes) [2, 29]. Given the information needs of contemporary IT organizations, SITBM requires a broader scope. Based on a three-part view of IT organizations' output (i.e. systems performance, information effectiveness, and service performance), Chang and King [8] propose an instrument to measure an IT organization's performance. These authors acknowledge, however, that the highly organization-specific structure of their instrument makes it unsuitable for external benchmarking. The approach of Fuerstenau [3] builds on a systematic analysis of an IT organization's capabilities by mapping these to governance, functional, process-related, and quality requirements. His instrument focuses on capturing organization-specific structures in the instrument. For SITBM, however, an organization-independent structuration of ITM is required.

The literature offers several approaches to structure ITM. They focus on IT applications [30], IT infrastructure [31], or strategic and organizational aspects [32]. Only a few approaches structure ITM as a whole. Earl [33], for instance, distinguishes between IS, IT, and information management strategy. Krcmar [34] uses a similar structuration, but adds an additional dimension which summarizes the executive functions of an IT organization (e.g. strategy, processes, governance). All of these approaches focus on an IT organization's internal value chain and only indirectly address interfaces to suppliers and customers. We argue that an IT organization's entire value chain and organizational environment should be explicitly reflected in an SITBM instrument. In this regard, Ward and Peppard [17] suggest structuring an IT organization along the entire value chain's various interfaces and requirements. Riempp et al. [35] have introduced a comprehensive reference framework (RF) that structures an IT organization into seven domains, three of which have interfaces to an IT organization's external environment (Figure 1).

The sourcing domain represents the IT organization's interface to the suppliers, while the service delivery domain represents the interface to the customers (i.e. business units). The four domains infrastructure, applications, processes, and projects represent the internal value chain. A seventh domain, comprising the IT strategy and governance areas, as well as financials and steering, unites these six domains. This last domain also represents the third interface which links to the IT organization's environment. In sum, Riempp et al. [35] capture an IT organization's entire value chain while their RF is sufficiently concise to support IT executives. Furthermore, it incorporates many IT professionals' practical experiences with IT organizations' requirements [40], which positively contributes to its usefulness. The RF does not predetermine any specific contents in the seven ITM domains. When the RF is used, it needs to be enriched with the actual contents required in a specific application scenario. This generic character of the ITM RF makes it well suited for SITBM where the specific contents are subject to ongoing change and refinement and need to reflect recent developments. In Section 6, we introduce a possible instantiation of the RF with contents as part of our expository instantiation: the RF itself is not part of our DT.

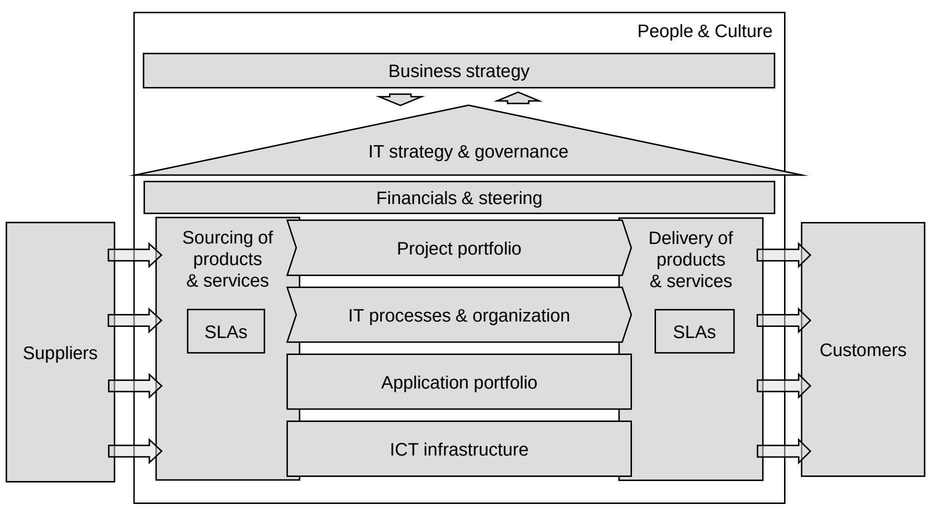

Figure 1. IT management reference framework [35]

2.2 Quality criteria applied in the design process

To determine the relevant quality criteria for SITBM methods, we searched the literature on benchmarking and IS design science and identified four quality criteria for benchmarking instruments (as part of our artifact on which we focus in this paper). Next, we introduce these quality criteria along with an initial set of DPs and propositions as part of our DT presented in Section 5. While the design principles capture important properties of an SITBM method, the propositions show how these properties are linked to the quality criteria. The initial DPs and propositions as derived from the literature guided our first iterations but were partially revised as we gained new empirical insights. The final version of our DT (including the final DPs and propositions) is presented in Section 5.

An SITBM instrument needs to cover the relevant ITM content (*content fit*) [8, 18, 36]. A sound strategic situation analysis requires an assessment of the most important IT services and a good understanding of the IT organization's processes and structures. Gimbert et al. [20] show that analyzing strategically irrelevant information reduces the results' validity, which does not contribute to improving an organization's current strategic position. Other researchers made similar observations [e.g. 7, 18]. Further, instruments used for IT strategic assessments should be kept as short as possible to be feasible for IT executives considering typical time and resource constraints [1, 8, 15]. Building on this, we derive our first DP:

DP1 *(Content fit tradeoff): The instrument should cover the relevant ITM facets to an extent manageable for IT executives.*

The relevant content has to be carefully quantified (*quantification quality*) to assure the benchmarking results' reliability [2, 21]. This means a benchmarking instrument needs to assure that the data collected from different companies can be compared. Comparability is assured by providing all the participants with precise definitions of the data to be collected. These definitions should be understandable and applicable, also in different organizational environments [11, 37]. Structuring single ITM domains (as provided by [30–32, 38]) is helpful in this regard because researchers dealing with specific ITM domains often provide comprehensive sets of appropriate PIs. SITBM mostly comprises qualitative and quantitative PIs. Definitional precision and a fit with established controlling and cost breakdown systems are key criteria for quantitative indicators [13]. Qualitative indicators are usually derived by analyzing the relevant facets of, for example, processes and deriving suitable questions on the state of implementation and maturity [2, 22]. To assure that the chosen PIs are appropriate for benchmarking, existing indicators should be used whenever possible[1]. While from a statistical perspective operationalizations should be as precise as possible, a tradeoff is needed in practice: a very detailed measure (e.g. cost breakdown) allows for precise peer group comparisons but is hardly manageable for the persons collecting the data [1, 11, 16]. Incorporating the participants' feedback is often the only way of determining the definitions' proper level of detail [16]. Hence:

***DP2** (Normalization): The instrument should employ normalizations of company-internal cost and quantity structures to facilitate a comparison with the peer group.*

The third quality criterion is contextualization, that is, the SITBM instrument's ability to collect context information. Context information has been shown to facilitate better interpretation of the results [7] as it helps understanding why certain results may have occurred [6, 37]. Some authors argue that many of the problems ascribed to unsuitable peer groups are actually problems of a lack of context information [14, 29]. The data collected with the SITBM instrument should thus be contextualized with organization-specific aspects, framing conditions, and additional strategic information. However, how this contextualization should be performed best in benchmarking

---

[1] See, for example, Chang and King [8] for meta-level operationalizations of qualitative measures regarding information effectiveness and systems and service performance; Fuerstenau [3] for qualitative and quantitative operationalizations of governance; and Khaiata and Zualkernan [5] or Fitoussi and Gurbaxani [39], for business-IT alignment or IT outsourcing measurement.

isnot suggested by previous research. While we acknowledge the principal relevance of contextualization, we defer the formulation of a corresponding DP to later stages of our research.

In summary, SITBM instruments fulfilling the quality criteria discussed so far have a positive impact on the quality of the SITBM results: content fit assures validity of the results, careful quantification contributes to reliability and comparability of the data, and contextualization increases validity and comparability. An improvement in the results' quality contributes to more substantial data analyses and, eventually, to the overall outcomes of the SITBM [6]. Thus, the better these three quality criteria are addressed, the higher is the utility of the SITBM instrument (i.e. enabling a comprehensive comparison of ITM). *Utility*, the fourth quality criterion we address, stems from IS design science research [27, 40, 41]. Venable [41] sees all design theories as utility theories that require "a particular type or class of technology (i.e. a meta-design), […] [to have] (some level of) utility (or usefulness) in solving or improving a problematic situation (with specified characteristics)" (p. 12). Hence, we propose:

P1a *(Utility): The better the instrument's content fit, the higher will be its utility.*

P1b *(Utility): The better the instrument's quantification quality, the higher will be its utility.*

P1c *(Utility): The better the instrument's contextualization, the higher will be its utility.*

The evaluation of an artifacts' utility is as crucial as the quality of the design although this evaluation is extremely complex [41]. Walls et al. [42] suggest measuring whether the meta-design satisfies the meta-requirements. Aier and Fischer [40] suggest that measures of utility should also account for the unintended side effects of using an artifact, since, for example, large negative side effects may significantly diminish its overall utility. In benchmarking, important criteria of utility are portability and privacy. Instruments should support distribution within one organization and portability across different organizations [11, 12, 43]. Data privacy is important since companies do not want to share their strategic data with competitors [7, 11, 28]. Given this background, we formulate:

DP3 *(Portability): The instrument should be easily portable and usable by different users in different environments.*

DP4 *(Privacy): The instrument should offer a high level of data privacy and allow for anonymization of the data of different participants.*

3. Research method

IS design research is concerned with the creation and evaluation of IT artifacts [44, 45], often together with IS design theories [26, 40, 42]. A DT is concerned with prescribing how a design process can be carried out [cf. 42, p. 37]. The primary addressees of IS design research are IS professionals, who demand knowledge on how to conduct IS initiatives and implement IS artifacts [46]. Accordingly, our approach to DT development builds on the conceptions of Gregor and Jones [26], who consider eight DT components that include both a theory-driven as well as an implementation-oriented perspective. The DT we propose is a theory for design and action [47] because it provides prescriptions on how to develop methods for SITBM. Complementary to this position, Baskerville and Pries-Heje [48] stress that "design theories (…) provide functional explanations of why designs and artifacts have certain attributes and features" (p. 273). Our DT explains the fundamental elements of any SITBM method, that is, how to design the data collection instrument (and corresponding glossary) and how to conduct an SITBM (i.e. role and process model). These abstractions into our DT are grounded in the design decision we took in each iteration and the corresponding evaluation of their impact on SITBM. This position allows us to establish a logical chain of evidence that supports "[…] inference[s] from the evaluation back to the design process and the conceptual understanding of the problem" [49, p. 828]. Such inferences are supported by different observations from our research process in which we show why the various architectural components included in our DT should be designed as described and formulate DPs and testable propositions to address the functional explanations. Finally, in line with recent IS design research [46], we hold that design theorizing is a socio-technical activity that has to explicitly consider the dynamic interactions between people and technology. We therefore opt for a research setting that includes multiple evaluation and theory development cycles to observe our artifact's use and evaluate our DPs in practice (Section 4) [40].

In our opinion, good design research theory should inform artifact development and new theoretical insights should be derived from artifact evaluation [27, 40]. Accordingly, our research approach is iterative: Over 10 years and five iterations during which we moved between rigor and relevance cycles [45], we developed the SITBM method that underlies the DT we propose here (i.e. the concrete artifact from which we draw inferences as to how SITBM methods need to be designed on a more general level). The first iteration started in 2004, and the last one finished at the end of 2014. Our research process per iteration builds on the suggestions by Peffers et al. [50]. Each iteration started with an analysis of the prior one which led to several design decisions that we complemented with new insights from research and practice. Based on these decisions, we revised our DT and discussed how the theory's components would be reflected in our method's design. Next, we applied our method in an SITBM. To assure that all participants would use the method consistently, we held introductory workshops during each SITBM round. After a round had been completed, we captured the participants' feedback in questionnaires and workshops. In addition, we reverted to our own observations to evaluate the design. For instance, we collected questions arising during the data collection phase and recorded misunderstandings and inconsistencies emerging during data analysis.

Table 1. Research methods applied per iteration

Iteration	Rigor cycle	Relevance cycle
0 – Prototype design cycle	• Literature search on benchmarking instrument design and ITM contents • Expert interviews on ITM contents	• Practitioner workshops to assess included contents, definitions and scales/measures • Pretest in seven companies
1 – Initial application	• Analysis and incorporation of results from pretest • Literature search to complement and integrate the findings	• Introductory workshops • (Qualitative) questionnaires and workshops to capture feedback • Researchers' observations
2 – Instrument redesign	• Qualitative analysis of feedback and observations (e.g. revision of definitions, measures, scales) • Literature search on ITM RFs to structure instrument • Practitioner feedback from previous iteration, interviews and workshops to determine relevant content after restructuration	• Introductory workshops • (Qualitative) questionnaires and workshops to capture feedback • Researchers' observations
3 – Incorporation of process model	• Qualitative analysis of feedback and observations (e.g. how to integrate the SITBM with the IT strategy process) • Literature search on contextualization of benchmarking data	• Case research involving interviews, document analyses, and observations • Data analysis using systematic open and selective coding
4 – Refinement	• Literature search on further contents to reflect recent developments in IT practice • Practitioner feedback from previous iteration	• Introductory workshops • (Qualitative) questionnaires and workshops to capture feedback • Researchers' observations
5 – Technical redesign	• Qualitative analysis of feedback and observations (e.g. revision of definitions, measures, scales) • Literature search on further contents to reflect recent developments in IT practice • Practitioner feedback from previous iteration	• Introductory workshops • (Qualitative) questionnaires and workshops to capture feedback • Researchers' observations

We used various qualitative research methods during the iterations (Table 1). In total, we applied our method 129 times to 102 companies in an SITBM initiative (27 companies participated more than once). The companies spanned 14 industries and ranged from very small to very big. They differed in their IT maturity, their approach to ITM, and the IT department's corporate role (e.g. innovator vs. support function).

4. Development process

Our research project started as a design science project aimed at developing an SITBM instrument. During iteration 3, however, we realized a stand-alone instrument was not sufficient for

SITBM, but that a process and role model would also be needed to guide the successful application of the instrument. This indicated that a more comprehensive method for SITBM was required. In the next sections, we describe how our research has progressed during the five iterations and delineate our lessons learnt from each iteration. In addition, we reflect on these lessons and derive our final set of propositions and DPs included in our DT. We strive to describe the DT as completely as possible since we intend our paper to also address a practical audience beyond the scientific community. We thus include propositions and DPs that might appear straightforward from an academic point of view but which we consider to be important in practice.

4.1. Prototype design cycle

The prototype design cycle aims at obtaining a feasibility prototype. Hence, our research process started with the development of an initial instrument for SITBM. In SITBM, instruments often take the form of questionnaires [43] which insure the comprehensive and anonymous collection of relevant data, two facets that are especially relevant in strategic contexts [7, 11, 28, 35]. These considerations resulted in a single questionnaire instantiated in Excel with 400 questions (statistical PIs, few qualitative assessments) focused on essential ITM issues such as basic costs and quantities, qualitative measures of IT strategy, and the corporate structure and demographics, which we derived from the practitioner and the scientific literature as well as from ongoing discussions with IS professionals. The latter approach has been suggested since in IS evaluation settings, the decision on the *relevant contents* to be benchmarked can often only be determined by practitioners [1, 2, 8]. As this is often overlooked by both researchers and practitioners [36], we extended our initial set of propositions to assure the completeness of our DT:

P2a *(Content fit): The higher the degree of practitioner input during the derivation of the PIs, the higher the content fit of the instrument.*

We pretested the instrument in seven companies where the users struggled with understanding the normalization rules and providing the correct data. For instance, some of the companies did not know how to handle part-time workers, semi-retired workers, or managers and group leaders responsible for various units. Consequently, we reformulated the problematic questions, compiled a glossary of the most important indicators and normalization rules, and, during our abstraction and reflection process of the prototype design cycle, introduced the following DP:

DP5 *(Glossary): The instrument should employ a glossary which includes definitions for the relevant PIs, their normalization rules, examples, and typical application scenarios.*

Later iterations showed that a glossary should be a mandatory component of an SITBM instrument which is why we included a glossary in the constructs of our DT. The need for a glossary has not yet been reported by benchmarking research [10].

4.2 First iteration – Initial application

We applied our instrument in two SITBMs in 2004 and 2005. We were unable to answer all questions posed by the 37 participants in the 2004 study on the definitions and normalizations because some questions were too detailed. This particularly pertained to aggregate cost measures, such as project costs. Many of the participants treated activities budgeted at less than EUR 20,000 not as projects to have less strict constraints and regulations. The inclusion of these low-cost activities in their project costs was complex or even impossible. We understood that a detailed definition of a project was necessary, as well as a scope statement and the establishment of a financial lower limit. We made similar observations with respect to other PIs. The relevance of the glossary as DT construct was underlined. Further, we learned that the mere existence of a glossary was not sufficient, but that precise prescriptions on the design of the glossary are necessary, especially regarding the level of detail of the glossary. Hence, we added the following proposition explaining the relevance of quantification quality:

***P3a** (Quantification quality): The more detailed the description and financial delineation of each PI, the higher will be the instrument's quantification quality.*

The quantification quality was also an issue in the final workshop of the 2004 SITBM. While most participants expressed their appreciation of the contents, they had concerns regarding the comparability of some indicators. We again revised the glossary and used the instrument in another SITBM with 41 companies in 2005. The feedback was encouraging as many of the participants maintained that the SITBM allowed them to compare their IT on a strategic level for the first time, thus indicating the utility of our method. The quantification quality seemed to increase, which is indicated by the fact that there were fewer questions in this regard.

4.3 Second iteration – Instrument redesign

After iteration 1, we had collected a vast number of suggestions on how to better capture the ITM content which we were unable to summarize in just a few changes. For example, many participants had asked for more detailed sections on sourcing, project management, and the respective core processes in these fields while others suggested capturing IT services (e.g. telephone, PCs, e-mail, archiving) in a more aggregated manner. As a result, we decided to split our instrument into several questionnaires on the basis of an existing ITM RF: we originally wanted to keep our instrument as parsimonious as possible and had hesitated to use a modular questionnaire design. From a functional perspective, by building on an existing RF we can better account for the demands for specialized sections while simultaneously increasing the participants' acceptance of the new structure. Several questionnaires also seemed more convenient to use because the large number of additions and changes made would have resulted in a significant increase in the instrument's length.

We conducted an extensive literature search for ITM RFs and, finally, structured our instrument according to the RF by Riempp et al. [35] (Figure 1) due to its empirical foundation and its coverage of an IT organization's entire value chain (see also Section 2.1). This structuring resulted in seven questionnaires. Next, we adjusted the PIs in the questionnaires using the participants' feedback and practitioner and scientific knowledge. For example, best-practice frameworks such as TOGAF were surveyed to identify relevant contents in the "IT applications" questionnaire (Section 6.3) and ITIL for the questionnaire "IT organization and IT processes" (Section 6.5). We included frameworks for enterprise architecture, project and risk management, and controlling. We also identified the core processes in the domains and included PIs to measure the state of these processes [e.g. 2, 3]. Further, we revised and extended the glossary (e.g. by adding a cost breakdown structure to illustrate the normalization rules) and added completion information on every question. To account for the longer instrument (1,000 items instead of 400), we defined one questionnaire as mandatory, the others could be filled out on individual interest. The resulting instrument was applied in an SITBM with 23 companies in 2007. Most of the participants appreciated the modularized structure: they could match their IT organization well with the instrument, easily distribute the questionnaires, and choose which questionnaires they would complete. The refined glossary and completion information also led to positive reactions.

During the abstraction and reflection process of this iteration, we ended up with several changes to our DT and new prescriptions. First, we adjusted our instrument construct so that it comprised several questionnaires for the domains to be captured in the SITBM, since the modularized structure had led to broad positive reactions concerning the convenience of use (easier distribution) and the content fit (individual choice of relevant contents). The individual choice of relevant contents also provided insights on artifact mutability; we were able to observe under which conditions participants chose to complete chosen questionnaires. The chosen RF was also appreciated by the participants, but most of them stated that *having* a modular structure based on an ITM RF improved the utility of the instrument, and that it was the coverage of ITM's value chain that contributed to a targeted choice of the relevant questionnaires. Given these insights, we adjusted DP1 (see Section 2.2) and added another DP along with three propositions:

DP1$_a$ *(Content fit tradeoff): The instrument should cover the relevant ITM facets in a modular structure that allows IT executives to flexibly adjust the contents to be benchmarked.*

DP6 *(Modular structure): The instrument should be structured using an ITM RF.*

P1d *(Utility): The more flexible IT executives can adjust the contents to be benchmarked, the higher will be the utility of the instrument.*

P2b *(Content fit): The better the RF used to structure the instrument covers an IT organization's value chain, the higher will be the content fit of the instrument.*

P3b *(Quantification quality): The more detailed the cost breakdown scheme for the normalization, the higher will be the quantification quality of the instrument.*

4.4 Third iteration – Incorporation of process model

Many comments and reactions in the second iteration pointed to the importance of better strategy process integration of the SITBM. In addition, SITBM projects seemed to be more successful if companies chose a participative approach to all the relevant stakeholders and if they provided more detailed guidelines on how to use the instrument. While our attention had been strongly focused on the contents and the quantification of the PIs, the participants' feedback highlighted the need to take a closer look at how the instrument was embedded and applied in different organizations. We also had not yet addressed our third quality criterion on contextualization. Consequently, in the third iteration round, we participated in SITBM projects at three companies. In this convenience sample, we observed our instrument's use during the IS strategizing process. The case studies took place in companies from different industries, of different sizes and structures from the summer of 2007 to the spring of 2009. Each case study took between two and five months to complete. For a detailed description of the cases and results see *(reference blinded)*.

The insights we generated during these case studies led to fundamental changes to our DT. We understood that a stand-alone instrument was not sufficient for SITBM, but that also a process and role model would be needed to guide the successful application of the instrument; that is, a method was needed. We, hence, decided to rescope our DT. While it was thus far intended to only provide guidance on the development of SITBM instruments and corresponding glossaries, we concluded it should also provide prescriptions for the development of respective process and role models. Our DT in its current form helps to develop entire SITBM methods. This change is reflected in the inclusion of two new constructs: process and role model. These constructs would ensure the instrument's successful application, the SITBM's proper integration into the overall strategic planning process, and the active involvement of ITM stakeholders. Concerning the quality criterion contextualization, we found that the stakeholders often took notes in the questionnaires to later jog their memories on how they made specific calculations or ended up with certain values. In addition, we collected strategic context information to better adjust the peer group. In all three cases, the notes and the additional information helped to better link the SITBM results back to the organization, improve their interpretability and acceptance. Hence, we added free-form comment fields in the questionnaires and a new section for capturing strategic context information. Concerning our DPs and propositions, we derived:

DP7 *(Commenting): The instrument should allow participants to leave comments per PI.*

P4a *(Contextualization): The more fine-grained the single PIs can be enriched with additional* *information, the better will be the contextualization of the instrument.*

P4b *(Contextualization): The more comprehensive the instrument captures an IT organization's strategic context, the better will be the contextualization of the instrument.*

4.5 Fourth iteration – Refinement

In iteration three, the change in IT organizations towards IT service providers became apparent. Thus, we restructured and extended our fifth questionnaire "IT organization and IT processes" to better account for IT's role as a service provider. To that end, we bundled our expertise with that of a second scientific IT benchmarking initiative experienced in IT service management. We incorporated process PIs according to ITIL, extended process and service management sections, and added a section on assessing the service desk as the most important interface to the business users. The resulting instrument was used in an SITBM with 12 companies in 2010 and 2011 (four companies participated twice). For the first time, almost no questions were raised regarding the breakdown of the IT costs with most of the participants accepting the definitions. Some questions arose related to the new and revised sections and PIs (e.g. the service desk). Concerning our DT, we found most propositions and DPs confirmed and did not introduce any adjustments.

4.6 Fifth iteration – Technical redesign

In this iteration, which started in 2011 and finished by the end of 2014, we did a technical redesign. During this iteration, 13 companies (three companies participated twice) used the instrument. While the participants' feedback on the content and definitions was only marginal, they did ask for a browser-based implementation instead of Excel. Although Excel had been the best solution to address portability, context mutability, and the security concerns of the participants for many years, today, web-based solutions seem more suited to implement the instrument: They fulfill the desired security standards similarly well, but are more easily portable and accessible from different devices. Still, the participants raised concerns whether a web-based solution would meet their individual data protection and privacy needs. Consequently, cloud-based solutions were rated as inappropriate and the instrument has been migrated into an online tool.

The feedback from the two most recent iterations leads us to conclude that our DT has gained a relatively high level of stability and applicability in terms of its content fit, quantification quality, contextualization, and utility: no new DPs, constructs, or propositions were identified in the last two iterations. Parts of the content have been adjusted to reflect new trends such as “bring your own device”, cloud and mobile computing in the sourcing domain (see Section 6.2).

5. A design theory for SITBM methods

Reflecting on the insights from our iterations, we end up with several changes to our initial design, DPs, and propositions as discussed in Sections 2 and 4.1. The underlying mechanisms and design choices that led to our final artifact form the basis of our DT (Table 2). An architectural blueprint of our design components can be found in Figure 2.

The *purpose and scope* of our DT is the development of SITBM methods that allow a comprehensive comparison of ITM with a peer group. The artifacts resulting from the DT will be applicable to any IT organization, but adjustments may be necessary (see artifact mutability).

Our DT comprises four sub-parts as *constructs* each SITBM method should contain: instrument (i.e. questionnaires), glossary, process model, and role model (Figure 2). As all sub-parts comprise an individual design, we will describe them only briefly here. Contrary to our initial design, an SITBM instrument should include single questionnaires to capture the ITM domains.

Table 2. Design theory for SITBM methods

Component	Description
Purpose and scope	Prescriptions for developing SITBM methods that allow for a comprehensive comparison between an IT organization's ITM and that of a peer group
Constructs	• Single questionnaires covering ITM domains based on an ITM RF • Glossary with definitions, examples, and typical application scenarios • Process model • Role model
Principles of form and function	DP1 and 6 address content fit, DP2 and 5 address quantification quality, DP3 and 4 address utility, and DP7 addresses contextualization
Artifact mutability	Each iteration provided insights into the mutability conditions of our DT • Contents, questionnaire structuring and composition, glossary and definitions may have to be adjusted for specific industries or organizational contexts • The heterogeneity of the peer group impacts the level of detail and normalization required in the instrument • The role and process models may require adjustment to a company's specific roles and process structures
Testable propositions	10 propositions (P1 through P4 with sub-propositions) were formulated to test the effects of different configurations of our DPs on the four quality criteria
Justificatory knowledge	DPs and propositions as well as relevant quantitative and qualitative indicators were derived from scientific and practitioner literature on ITM and benchmarking. In addition, data and observations from the iterations helped to revise and extend our initial DPs and propositions
Principles of implementation	• Identify relevant content based on the continuous analysis of the literature and on practitioner-driven discussions • Pretest the questionnaire in several companies and in different infrastructural environments to assure content fit and universal understanding of the definitions • Use a technical basis that allows for the easy adaption of the structure and contents • Conduct introductory workshops at the beginning of an SITBM to assure that all the participants understand the instrument
Expository instantiation	• A Microsoft Excel-based questionnaire comprising seven sub-questionnaires (iterations 1 through 4) • A web-based SITBM questionnaire (since iteration 5)

Single questionnaires allow for distributed data collection and help IT executives better link the results to different organizational domains. A detailed glossary containing a normalized cost breakdown structure and precise PI definitions is needed to ensure that participants have the ability to transform their internal structures into a normalized form that can be used for comparison. A process and role model assure sound embedding of the SITBM in an IT organization's strategy process. The role model comprises the relevant IT stakeholders to be involved in the SITBM as well as their roles and responsibilities during the entire SITBM process. The SITBM process steps are described in the process model (Figure 2, simplified illustration based on Camp [11]). Similar to other researchers [e.g. 10, 11, 51], we had initially assumed that the employed process model should be considered independently from the instrument. However, in our third iteration, we found that the process model should include prescriptions on how to collect and contextualize the data using a specific instrument. While a few other benchmarking researchers have also used this approach [43, 51], it does not yet seem to have found broader acceptance [36]. Since the focus of this paper is on the instrument, we do not describe the process and role model and their design in more detail here.

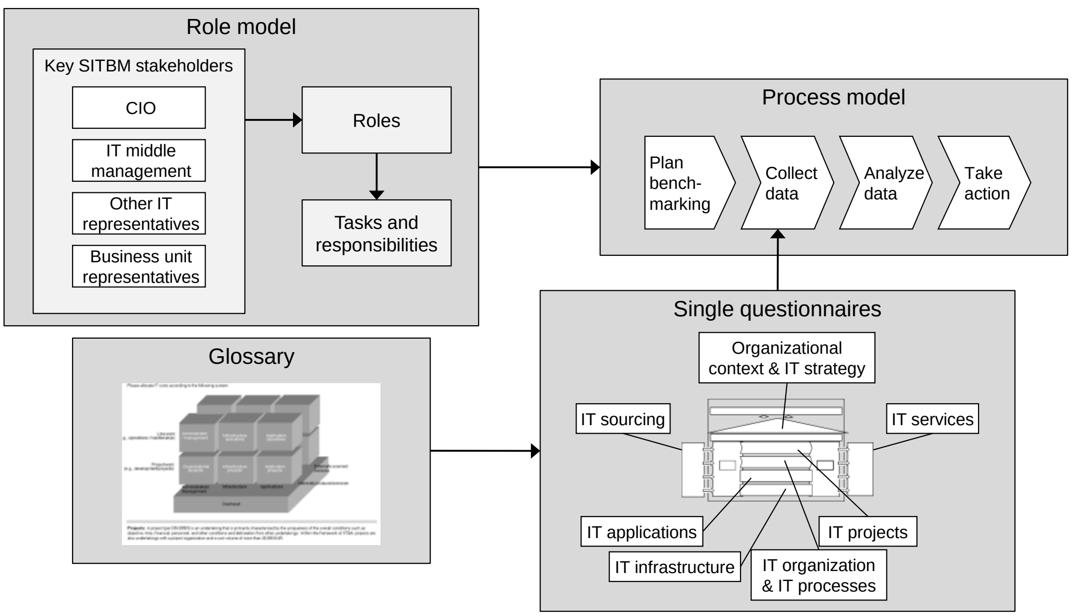

Figure 2. Architectural blueprint of the design components

The constructs are closely linked to the *principles of form and function* that provide an "abstract 'blueprint' or architecture that describes the artifact" [26, p. 323]; in our case for the construction of SITBM methods. The principles address our quality criteria and increase the validity, reliability, and interpretability of the SITBM results. For example, the participants reacted positively to exemplary calculations and application scenarios in the glossary since these helped them quickly understand the often complex definitions.

The application of our artifact in 102 different organizational settings offered valuable insights towards *artifact mutability*. We could observe how our instrument reacted to changing use conditions. For example, the contents and the composition of various questionnaires may have to be adjusted for specific organizational contexts. Companies of small sizes, specific industries, or legal structures may benefit from a different structuring of the instrument. For instance, according to our observations, companies with high outsourcing quotas were mostly interested in sourcing and project management. Often, such companies requested more depth regarding these aspects and opted for reducing the parts on infrastructure and applications. In addition, a group of energy providers adjusted various definitions and the cost breakdown to their joint understanding. The role and process models may need adjustment to a company's specific roles and process structures. In heterogeneous peer groups, the normalization rules seem to require more detail than in companies of similar industries or strategic orientation.

Our efforts yielded 10 *testable propositions*, which correspond to our quality criteria and are based either on the justificatory knowledge or our insights gained in the iterations. These propositions highlight how instantiating the DPs of our proposed design theory in any SITBM method is linked to that method's quality and performance. Accordingly, these propositions are helpful in developing test cases and evaluating future instantiations of the DT, facilitating a high quality of future artifact instantiations. The DT, in turn, will have to prove that such future instantiations contribute to a structured comparison of ITM. Testing whether the propositions holds or not in future instantiations, thus, will also contribute to the extension and revision of the design theory we propose here.

According to Gregor and Hevner [27] *justificatory knowledge* includes "any knowledge that informs design research, including informal knowledge from the field and the experience of practitioners" (p. 340). Hence, we have also included the feedback of participating organizations as far as the feedback could be isolated from a specific organizational context. For instance, we analyzed the collected materials for emerging patterns and clustered them into problem types (e.g. problems related to missing content, definitions, or technical bugs). For these problem types, we discussed which of the problems may be generally relevant for our design and which are rather specific and, therefore, provide insights on artifact mutability. Our perspective was enriched by a comparison with the literature. We searched the scientific and the practitioner literature on ITM and benchmarking to identify relevant measures and derive respective PIs. Our initial DPs and propositions (Section 2) were also derived from this literature.

We have identified four *implementation principles*. First, the contents covered should be subject to continuous revision and relevance checks. While the literature provides a sound basis for SITBM contents, the instrument should specifically reflect the topics of current relevance to IT executives to ensure that recent developments and their impact can be addressed. Second, the instrument should be pretested in several companies and in different infrastructural environments to assure its content validity and the clarity of definitions and PIs. Third, the chosen technological basis should allow an easy modification of the structure and the content because constant instrument change is a core SITBM characteristic. In our setting, a spreadsheet-based solution was a suitable technological basis for long time. However, modern content management systems and online survey tools allow for a similar degree of adaptability and may better address contemporary IT executives' needs regarding usability and mobility. Finally, we found that introductory workshops at the beginning of an SITBM are valuable to assure that all participants understand the instrument and will apply it correctly.

6. Expository instantiation

In this section, we introduce the instrument as the central part of our expository instantiation. The current version of the instrument comprises seven sub-questionnaires covering the ITM domains according to Riempp et al. [35]. These domains are instantiated with contents as shown in Table 3. This instantiation should not be viewed as complete or universal but rather as a suggestion and starting point for further refinement in a specific practical setting (e.g. industry-specific adjustments). We collect 1,608 quantitative and qualitative data points about one IT organization. Quantitative measures are used for costs and other countable facts, while qualitative measures collect “softer” and process-related facts, using nominal or Likert scales.

We have also incorporated comment fields for context information and detailed definitions of each item. On the basis of the collected information, we calculate a set of 130 PIs (Table 3) such as the IT costs per turnover, the margin, or the outsourcing quota, but also subject-specific indicators such as the average ticket backlog per service desk employee. The qualitative data is used to derive simple maturity indicators (with values from 0% to 100%) for specific process domains (e.g. the IT strategy process or IT service management), whereby 0% indicates that the respective process does not exist and 100% indicates that all process facets exist in the organization. Each process domain is measured using 6 to 18 items which are condensed using the arithmetic mean scaled to 100%.

Table 3. Structure of our SITBM instrument

Questionnaire	ITM Domain	Topics	Examples of PIs
Organizational context & IT strategy	Business strategy, IT strategy & governance, financials & steering	• Organizational classification • General company data • IT strategy • Hot IT topics • Cost and quantity structures	• IT costs per turnover / employee • Outsourcing quota • # of external per internal employees • Project percentage of total IT costs • Maturity of IT strategy process
IT sourcing	Suppliers; sourcing of products & services	• IT sourcing strategy • Management of IT sourcing • Spectrum of sourced IT ser-	• Maturity of the IT sourcing strategy process • Maturity of the management of exter-

		vices • Success/costs of IT sourcing	nally sourced services • Motivation for sourcing
IT applications	Application portfolio	• Application portfolio • Costs and quantities of IT applications • State of the application and architecture management • Complexity of application landscape	• Quota of standard applications • # of applications per IT-supported employee • Maturity of application and architecture management • Complexity factor / costs of application software
IT infrastructure	ICT infrastructure	• Costs/quantities of IT infrastructure • Complexity/costs of infrastructure (software) landscape	• Ratio of infrastructure software to hardware costs • Cost quota servers/notebooks • Complexity factor notebooks/ infrastructure software
IT organization & IT processes	IT processes and organization	• IT service management and processes • Reporting, budgeting, and controlling processes • Service desk	• Maturity of IT service management • Maturity of IT processes • Average backlog per service desk employee • Average service desk cost per ticket
IT projects	Project portfolio	• Strategic and operative project management • Process models/frameworks • Project success	• Percentage of projects in time /budget/quality/cancelled projects • Maturity of strategic/operative project management
IT services	Delivery of products & services; customers	• Structure of service catalogue • Costs and performance of the most important IT services	• Costs of the provision of a laptop • Costs and size per mailbox • Costs per terabyte of storage • Frequency of backups / archiving

6.1 Organizational context and IT strategy

Besides collecting information about the current IT strategy, the questionnaire is particularly relevant for contextualizing SITBM results by capturing data on a company's industry, size, and turnover, its strategic orientation (i.e. enabler vs. supporter, growth vs. downsizing), the standardization of the processes, cost structures, etc. These characteristics help a participant company better determine a peer group and foster the interpretability of the SITBM results. A cost breakdown is done of the overall IT costs and quantities with respect to, for example, the sourcing, line work, projects, and external and internal IT employees. Another part of the questionnaire captures information about the IT organization's past, current, and future project portfolio. To this end, we provide an annually updated list of "hot IT topics" that are either included in trend reports, or extensively discussed in practitioner literature. We measure the standardization and maturity of the IT strategy process with 13 items covering businesses' IT alignment, handling of strategic goals, IT architecture, and application portfolio. Examples of these items are:

> *The IT strategy is derived from the business strategy and supports it efficiently and effectively./ The IT's strategic goals are realized and controlled via established implementation processes./ The implementation of the IT architecture process is anchored in the organization via a staff position or a dedicated work group.*

We use 4-point Likert scales to measure the as-is state and 5-point Likert scales to capture an item's perceived importance. The importance of an item does not influence the maturity indicator but does support later interpretation and helps the weighting of the improvement fields (e.g. certain facets of an IT strategy process are more important for an enabler than for a supporter [33]). Each question has a detailed definition to assure a high degree of comparability and to cope with confirmation bias. For instance, the definition of the abovementioned item 2 is as follows:

> *"Is completely true" means that processes established in the organization are used to realize and control the attainment of strategic IT goals. "Is rather true" reflects the processes' lower level of maturity. The additional gradations reflect that these processes are only partially used for strategic IT goals.*

This first questionnaire is by far the longest, as well as the most fundamental questionnaire in our instrument, as it captures all data relevant to identify the peer group and basic strategic position. Hence, it is mandatory for all participants while the others are optional.

6.2 IT sourcing

IT sourcing is of strategic interest, especially in light of IT industrialization and efficient service delivery [e.g. 39]. Thus, this questionnaire investigates the services and processes that IT organizations source externally and why they do so. We also account for the success and costs of outsourcing. The questionnaire collects data on the IT sourcing strategy and the management of sourcing (e.g. an ongoing comparison of the internal and the external service costs, the use of structured processes to decide whether services should be delivered internally or externally, or the steering of externally sourced services). Underlying items measure, for example, the degree to which underpinning contracts or incentive systems exist. Overall, the questionnaire supports IT executives as they can judge and improve the contents of their sourcing strategy by analyzing the behaviors of their peers.

6.3 IT applications

The development, operation, and maintenance of IT applications are some of the key cost and complexity drivers in an IT organization. We capture the costs, size, and efficiency of a company's application portfolio as well as its respective management processes. We account for the degree of standardization and the usage of application frameworks as important levers for cost savings. Because software licenses are another IT cost driver, we investigate the license management's maturity. As a mature enterprise architecture management is crucial for assuring the business-IT alignment and compliance, we capture the actual existence of an enterprise architecture management along with its integration into long-term IT planning, whether committees and organizational units control the architecture's effective implementation and standard conformity, and whether efficient control mechanisms are employed to monitor the status of the applications.

6.4 IT infrastructure

While IT infrastructure is no longer a competitive differentiator for most companies, it is still considered the backbone of a company and its availability remains a key prerequisite for and enabler of higher-order IT effects. Hence, shedding light on problematic infrastructure components is an important ITM facet. Given the new trends toward in-memory computing, cloud services, and Big Data analytics, this field is regaining strategic relevance. While we currently analyze the IT infrastructure landscape's complexity and costs, as well as its current level of standardization and virtualization, we will soon have to extend the questionnaire with sections including the abovementioned trends to capture the relevant ITM content.

6.5 IT organization and IT processes

This questionnaire reflects IT organizations' role as service providers. The first part of the questionnaire determines the maturity of IT service management; that is, for instance, how service and operational level agreements are implemented and monitored. We also account for adopted process frameworks such as ITIL, CobiT, or ISO20000, and for the implementation status of typical IT processes, such as incident management, change management, service level management, demand management, and service desk. With regard to the service desk, we also capture detailed information about its costs, personnel, and ticket quantities. In doing so, we meet the demands of many IT executives who requested a detailed analysis of this important interface to the business. The second part of the questionnaire measures the maturity of the reporting and budgeting processes. In summary, benchmarking the IT service processes and organizational structures provides IT executives with valuable insights into how to improve their service delivery and customer satisfaction.

6.6 IT projects

Although the majority of IT investments are implemented through projects, many IT projects do not meet their objectives in terms of budget, time, and quality. IT projects are, therefore, one of the most important components of an IT organization's internal value chain and must be analyzed in an SITBM. We investigate the state of project success by collecting information on the number of projects meeting budget, time, and quality conditions as well as the number of cancelled projects. The instrument investigates the project portfolio management (e.g. the portfolio definition, resource allocation processes, project risk management, and beneficial ITM routines) and the operational project management (e.g. career or compensation models for project managers, charging models for project costs, and resource request handling before and after the project kick-off) in detail. These issues are complemented with an analysis of the adopted project management methodology. Quantitative project data, such as the costs, number of project employees, etc., are collected in the questionnaire “Organizational Context and IT Strategy.”

6.7 IT services

We analyze the most important delivered IT services on a relatively high abstraction level by examining the aggregated costs or prices as well as the overall service performance. An SITBM is not price benchmarking and can thus hardly provide the same level of price development detail. However, when revising the IT strategy, or deriving long-term IT initiatives based on the SITBM, IT executives benefit in a limited way from statements indicating that, for instance, a printed page costs 30 cents more than it does at peer companies. To this end, the questionnaire focuses on the nine most often offered IT services (PCs and laptops, servers, database systems, e-mail, archiving, back-up, SAP, telephony and network services, and storage), and sheds light on their specific high-level PIs. Examples of such indicators are the costs of the provision of a laptop, the costs and size per mailbox, the costs per terabyte of storage, and the frequency of back-ups and archiving. We also capture general structural data on the service catalogue, such as the number of IT end-users, the size of the service catalogue, and the number of agreed on service levels. To sum up, this questionnaire completes the SITBM by analyzing an IT organization’s delivered products and services on an aggregation level suitable for IT executives.

7. Discussion and Conclusion

A DT comprises knowledge of artifact development, its use, and the artifact itself. According to the Knowledge Contribution Framework [27], our research belongs to the “improvement” category: while there is an abundance of knowledge on benchmarking and IT management, solutions for benchmarking instruments are scarce. The most comprehensive instrument to assess the performance of the IS function is introduced by Chang and King [8]. Other works focus on an assessment of the enterprise architecture and governance [3, 5] or on the benchmarking of IT-supported processes [2]. Our research extends these previous approaches by providing a comprehensive DT for SITBM methods that includes not only an instrument, but also a glossary, a process model, and a role model. Our 10 years of research allowed us to improve our DT and SITBM method while gaining in-depth insights as to why our design and implementation principles work the way they do. The inclusion of a process and role model into our DT has proven to be crucial to guide the successful application of the instrument. Some of our DPs even allowed for before/after comparisons. For instance, DP1 on content fit tradeoff was adjusted after we had evaluated our initial design in various organizational settings. In these settings, a modular and distributable structure was superior to a condensed and compact instrument structure. Covering more content in the instrument (i.e. a large number of collected data points) resulted in higher utility perceptions of the practitioners as long as a modular structure allowed flexible and easy adjustment of the contents to be benchmarked. Both findings are opposed to previous recommendations made for benchmarking instrument development [1, 8, 15].

We also present an expository instantiation of the instrument part of our DT. Structured and personal feedback was continuously captured to evaluate the instrument's utility. The participants were asked how relevant the measured contents were for an assessment of ITM, how well they could match their internal structures, cost breakdown logic and definitions with the instrument and employed ITM RF, and, given that the employed ITM framework is from 2008, whether it is dated and should be replaced. The participants in each round stressed the importance of having a RF and did not articulate any strong objections to the existing RF. Still, the finding we have anchored in our DT is that using any RF to modularly structure ITM will positively contribute to the quality of the overall instrument, and thus, the SITBM method (e.g. DP1, DP6 or P2b). The choice of a specific ITM RF and its instantiation with contents depends on the given application scenario. Given this background, practitioners benefit from our instrument by learning about the ITM domains they should capture in their strategic assessments. We also provide them with validated PIs. Our long-term engagement also helped rule out too imprecise or technical definitions – an issue often raised by practitioners [8, 15, 22]. Our work provides practitioners with guidelines on how to develop and implement own SITBM methods and how to conduct the whole SITBM project. We especially highlight that structure and contents of an SITBM instrument are separate entities which might need to be adjusted in a given application setting. Researchers can draw on our insights into the contents and processes when considering the quality criteria of ITM and SITBM instruments and how these criteria can be addressed during instrument development and usage. We also underline some of SITBM's key requirements, such as integration into the overall strategy process and proper contextualization. Many of the companies involved in our research over the years reported that the SITBM method we used, which followed our DT, helped them to revise their IT strategy and identify areas of future improvement. We thus contribute to practice and research by reporting a field-tested SITBM instrument and providing a DT for developing SITBM methods.

We acknowledge several limitations of our research. First, our SITBM participants constitute a convenience sample: all of them participated in the same SITBM initiative. Future research may need to evaluate our DT and any artifact instantiated based on it in the context of other initiatives. Second, we opted for a close cooperation with the industry from the beginning: we developed an artifact for IT managers. By summarizing and scientifically reflecting on what we learnt from this process we were able to derive the proposed DT. While this approach presents potential, such as high relevance and utility, other, more theoretically driven approaches to instrument development may identify other useful instruments for SITBM after fewer iterations. Third, the chosen ITM RF of Riempp et al. may have impacted some of the design decisions. Using a different ITM RF may have led to different observations. Still, we believe our main findings would have been preserved since these are independent from any given RF.

Improvement is a key objective of our research. Hence, future researchers may also want to compare the utility of our DT and method with those of other approaches to SITBM, such as consortia or partner benchmarkings [40, 41]. We expect to conduct future development rounds because ITM's relevant content, structure, and interfaces are likely to develop further in the future. Finally, future research could also investigate in more detail how different companies use our method. As we observed, some of our participants primarily used the indicators and structures in our instrument to improve their IT controlling, but not to compare themselves to others. Such unintended side effects may have positive or negative consequences [41, 44] and thus inspire further investigation.

8. References

[1] A. Alshawaf, J. Ali, and M. Hasan, "A benchmarking framework for information systems management issues in Kuwait," *Benchmarking: An International Journal*, vol. 12, no. 1, pp. 30–44, 2005.

[2] P. B. Cragg, "Benchmarking information technology practices in small firms," *European Journal of Information Systems*, vol. 11, no. 4, pp. 267–282, 2002.

[3] D. Fuerstenau, "Reframing the Governance Debate: A Multilevel Performance Measurement Appraoch Based on Capabilities," *ECIS 2012 Proceedings*, 2012.

[4] M. C. Lacity and R. Hirschheim, "Benchmarking as a strategy for managing conflicting stakeholder perceptions of information systems," *The Journal of Strategic Information Systems*, vol. 4, no. 2, pp. 165–185, 1995.

[5] M. Khaiata and I. A. Zualkernan, "A Simple Instrument to Measure IT-Business Alignment Maturity," *Information Systems Management*, vol. 26, no. 2, pp. 138–152, 2009.

[6] B. Müller, F. Ahlemann, and K. Roeder, "Understanding the success of strategic IT/IS benchmarking: results from a multiple-case study," in *Proceedings of the 18th European Conference on Information Systems (ECIS 2010)*, Pretoria, South Africa, 2010.

[7] H. A. Smith and J. D. McKeen, “Measuring IS: how does your organization rate?,” *SIG-MIS Database*, vol. 27, no. 1, pp. 18–30, 1996.
[8] J. C. Chang and W. R. King, “Measuring the Performance of Information Systems: A Functional Scorecard,” *Journal of Management Information Systems*, vol. 22, no. 1, pp. 85–115, 2005.
[9] J. Williams, C. Brown, and A. Springer, “Overcoming benchmarking reluctance: a literature review,” *Benchmarking: An International Journal*, vol. 19, no. 2, pp. 255–276, 2012.
[10] G. Anand and R. Kodali, “Benchmarking the benchmarking models,” *Benchmarking: An International Journal*, vol. 15, no. 3, pp. 257–291, 2008.
[11] R. C. Camp, *Benchmarking: The Search for Industry Best Practices That Lead to Superior Performance*. Milwaukee: Quality Press, 1989.
[12] G. H. Watson, *Strategic Benchmarking: How to Rate Your Company's Performance against the World's Best*. New York, NY: John Wiley & Sons, 1993.
[13] J. Maleyeff, “Benchmarking performance indices: pitfalls and solution,” *Benchmarking: An International Journal*, vol. 10, no. 1, pp. 9–28, 2003.
[14] G. Panagiotou, “Reference theory: strategic groups and competitive benchmarking,” *Management Decision*, vol. 45, no. 10, pp. 1595–1621, 2007.
[15] R. Ramabadran, J. Dean, J. Evans, and A. Raturi, “Testing the Relationship Between Team and Partner Characteristics and Cooperative Benchmarking Outcomes,” *IEEE Trans. Eng. Manage*, vol. 51, no. 2, pp. 208–225, 2004.
[16] D. Elnathan, T. W. Lin, and M. S. Young, “Benchmarking and Management Accounting: A Framework for Research,” *Journal of Management Accounting Research*, vol. 8, pp. 37–54, 1996.
[17] J. Ward and J. Peppard, *Strategic Planning for Information Systems*, 3rd ed. Chichester: Wiley, 2003.
[18] H. Newkirk, A. L. Lederer, and C. Srinivasan, “Strategic information systems planning: too little or too much?,” *Journal of Strategic Information Systems*, vol. 12, no. 3, pp. 201–228, 2003.
[19] A. L. Lederer and V. Sethi, “Key prescriptions for strategic information systems planning,” *Journal of Management Information Systems*, vol. 13, no. 1, pp. 35–62, 1996.
[20] X. Gimbert, J. Bisbe, and X. Mendoza, “The role of performance measurement systems in strategy formulation processes,” *Long Range Planning*, vol. 43, no. 4, pp. 477–497, 2010.
[21] L. C. Carpinetti and A. M. de Melo, “What to benchmark?: A systematic approach and cases,” *Benchmarking: An International Journal*, vol. 9, no. 3, pp. 325–441, 2002.
[22] A. del-Río-Ortega, M. Resinas, C. Cabanillas, and A. Ruiz-Cortés, “On the definition and design-time analysis of process performance indicators,” *Special section on BPM 2011 conference*, vol. 38, no. 4, pp. 470–490, 2013.
[23] R. P. Bagozzi, “Measurement and Meaning in Information Systems and Organizational Research: Methodological and Philosophical Foundations,” *MIS Quarterly*, vol. 35, no. 2, pp. 261–292, 2011.
[24] S. MacKenzie, P. Podsakoff, and N. Podsakoff, “Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques,” *MIS Quarterly*, vol. 35, no. 2, pp. 293–334, 2011.
[25] V. Serafeimides and S. Smithson, “Information systems evaluation in practice: a case study of organizational change,” *Journal of Information Technology*, vol. 15, no. 2, pp. 93–105, 2000.

[26] S. Gregor and D. Jones, “The anatomy of a design theory,” *Journal of the Association of Information Systems*, pp. 312–335, 2007.
[27] S. Gregor and A. R. Hevner, “Positioning and presenting design science research for maximum impact,” *MIS Quarterly*, vol. 37, no. 2, pp. 337–356, 2013.
[28] C. S. Saunders and J. W. Jones, “Measuring Performance of the Information Systems Function,” *Journal of Management Information Systems*, vol. 8, no. 4, pp. 63–82, 1992.
[29] T. Clayton and B. Luchs, “Strategic benchmarking at ICI Fibres,” *Long Range Planning*, vol. 27, no. 3, pp. 54–63, 1994.
[30] A. H. Segars and V. Grover, “Strategic Information Systems Planning Success: An Investigation of the Construct and Its Measurement,” *MIS Quarterly*, vol. 22, no. 2, pp. 139–163, 1998.
[31] M. Mocker and A. Teubner, “Towards a Comprehensive Model of Information Strategy,” in *Proceedings of the Thirteenth European Conference on Information Systems (ECIS 2005)*, 2005, pp. 747–760.
[32] D. Boddy, A. Boonstra, and G. Kennedy, *Managing Information Systems: An Organisational Perspective*, 2nd ed. Harlow, UK: Pearson, 2005.
[33] M. J. Earl, “Integrating IS and the Organization: A Framework of Organizational Fit,” in *Information Management: The Organizational Dimension*, M. J. Earl, Ed, Oxford: Oxford University, 1996.
[34] H. Krcmar, *Informationsmanagement*, 5th ed. Heidelberg: Springer, 2009.
[35] G. Riempp, B. Müller, and F. Ahlemann, “Towards a Framework to Structure and Assess Strategic IT/IS Management,” in *Proceeding of the 16th European Conference on Information Systems (ECIS 2008)*, 2008, pp. 2484–2495.
[36] P. Hong, S. W. Hong, J. J. Roh, and K. Park, “Evolving benchmarking practices: a review for research perspectives,” *Benchmarking: An International Journal*, vol. 19, no. 4/5, pp. 444–462, 2012.
[37] M. Hinton, G. Francis, and J. Holloway, “Best practice benchmarking in the UK,” *Benchmarking: An International Journal*, vol. 7, no. 1, pp. 52–61, 2000.
[38] M. T. Smits, K. G. van der Poel, and P. M. A. Ribbers, “Assessment of information strategies in insurance companies,” in *Strategic Information Management: Challenges and strategies in managing information systems*, R. D. Galliers and D. E. Leidner, Eds. 3rd ed, Oxford: Butterworth-Heinemann, 2003, pp. 64–88.
[39] D. Fitoussi and V. Gurbaxani, “IT Outsourcing Contracts and Performance Measurement,” *Information Systems Research*, vol. 23, no. 1, pp. 129–143, 2012.
[40] S. Aier and C. Fischer, “Criteria of progress for information systems design theories,” *Inf Syst E-Bus Manage*, vol. 9, no. 1, pp. 133–172, 2011.
[41] J. R. Venable, “The Role of Theory and Theorising in Design Science Research,” in *Proceeding of the 1st International Conference on Design Science Research in Information Systems and Technology (DESRIST 2006)*, 2006, pp. 1–18.
[42] J. G. Walls, G. R. Widmeyer, and O. A. El Sawy, “Building an Information System Design Theory for Vigilant EIS,” *Information Systems Research*, vol. 3, no. 1, pp. 36–59, 1992.
[43] H.-L. Chen, “A competence-based strategic management model factoring in key success factors and benchmarking,” *Benchmarking: An International Journal*, vol. 12, no. 4, pp. 364–382, 2005.
[44] S. T. March and G. F. Smith, “Design and natural science research on information technology,” *Decision Support Systems*, vol. 15, no. 4, pp. 251–266, 1995.

[45] A. R. Hevner and S. Chatterjee, *Design Research in Information Systems*. New York, Heidelberg: Springer, 2010.
[46] S. A. Carlsson, S. Henningsson, S. Hrastinski, and C. Keller, “Socio-technical IS design science research: developing design theory for IS integration management,” *Inf Syst E-Bus Manage*, vol. 9, no. 1, pp. 109–131, 2011.
[47] S. Gregor, “The Nature of Theory in Information Systems,” *MIS Quarterly*, vol. 30, no. 3, pp. 611–642, 2006.
[48] R. Baskerville and J. Pries-Heje, “Explanatory Design Theory,” *Bus Inf Syst Eng*, vol. 2, no. 5, pp. 271–282, 2010.
[49] H. Meth, B. Mueller, and A. Maedche, “Designing a Requirement Mining System,” *Journal of the Association for Information Systems*, vol. 16, no. 9, pp. 799–837, 2015.
[50] K. Peffers, T. Tuunanen, M. Rothenberger, and S. Chatterjee, “A Design Science Research Methodology for Information Systems Research,” *J. Manage. Inf. Syst*, vol. 24, no. 3, pp. 45–77, 2007.
[51] S. Drew, “From knowledge to action: the impact of benchmarking on organizational performance,” *Long Range Planning*, vol. 30, no. 3, pp. 427–441, 1997.