Balanced Scorecard for Method Improvement: Approach and Experiences

Kurt Sandkuhl[1,2], Ulf Seigerroth[2]

[1]University of Rostock, Albert-Einstein-Str. 22, 18059 Rostock, Germany
`kurt.sandkuhl@uni-rostock.de`
[2]Jönköping University, School of Engineering, Box 1026, 55111 Jönköping, Sweden
`{kurt.sandkuhl, ulf.seigerroth}@ju.se`

Abstract. Modelling methods provide structured guidance for performing complex modelling tasks including procedures to be performed, concepts to focus on, visual representations, tools and cooperation principles. Development of methods is an expensive process which usually involves many stakeholders and results in various method iterations. This paper aims at contributing to the field of method improvement by proposing a balanced scorecard based approach and reporting on experiences from developing and using it in the context of a method for information demand analysis. The main contributions of the paper are (1) a description of the process for developing a scorecard for method improvement, (2) the scorecard as such (as a tool) for improving a specific method, and (3) experiences from applying the scorecard in industrial settings.

Keywords: Method improvement, balanced scorecard, method engineering, information demand analysis method

1 Introduction

Modeling methods provide structured guidance for performing complex modeling tasks including procedures to be performed, concepts to focus on, visual representations for capturing modeling results, tools and cooperation principles (see section 2.1). Method engineering (ME) is an expensive and knowledge intensive process, which usually involves many stakeholders and results in various engineering iterations. This paper aims at contributing to the field of method engineering and especially method improvement by proposing a balanced scorecard (BSC) based approach and reporting on experiences from using the method and the BSC in the context of a method for information demand analysis.

The primary perspective for method improvement taken is that of an organization using the method for business purposes and aiming at improving the contribution of the method to business objectives. In this context, approaches from the field of "business value of information technology" (BVIT) are relevant and were investigated. Most of the BVIT approaches currently existing originated in a demand from enterprises to evaluate the contribution of IT to the business success. Section 2.2 includes an overview to BVIT approaches.

One of the general approaches for measuring BVIT is to capture indicators for different perspectives of the business value in a scorecard with different, balanced perspectives. Our proposal is to apply this approach for method improvement as it allows for combining different aspects relevant for business value, such as quality of the results achieved by using the method, quality of the method documentation and quality of the work procedures included in the method. The application of the scorecard is illustrated using the method for information demand analysis (the IDA-method). For the organizations using the IDA-method in own projects, the scorecard was supposed to be a management instrument for the operational use of the method.

Main contributions of the paper are (1) a description of the process for developing a scorecard for method improvement, (2) the scorecard as such (as a tool) for improving the IDA-method, and (3) experiences from applying the scorecard in industrial settings. The remainder of this paper is structured as follows: Section 2 summarizes the foundation for our work from method engineering and business value of IT research. Section 3 introduces the research approach taken. Section 4 describes the development process of the BSC and the resulting "method scorecard". Section 5 is dedicated to experiences related to the use of the scorecard. Section 6 summarizes the results and gives an outlook on future work.

2 Theoretical foundation

2.1 The Notion of Methods

Methods are often used as instrumental support for development of enterprises, e.g. for enterprise modeling (EM), information systems design etc. According to our view the use of methods is to be regarded as artifact-mediated actions where different prescribed method actions will guide our development work. A method as an artifact is something that is created by humans and the artifact can't exist without human involvement either by design or by interpretation (c.f. [11, 9]). An artifact can therefore be instantiated as something with physical- and/or social properties which also needs to be taken into consideration during method improvement. Method-mediated actions are diverse in nature and can also be tacit in character. Guidance for actions can also be found in the solution space in terms of artifacts like best practices, which also can be instantiated through different patterns. Methods are in many cases also implemented in computerized tools to facilitate the modeling process.

Our focus is on methods, method engineering, and method improvement. We acknowledge the ISO/IEC 24744:2014 standard and this definition of methods. During modeling activities, there is usually a need to document various aspects, and many methods include rules for representation, usually referred to as modeling techniques or notation rules. Methods also provide procedural guidelines (work procedures), which often are tightly coupled to notation. The work procedures involve meta-concepts such as process, activity, information, and object, as parts of the prescribed actions. The work procedures are also parts of the semantics of the notation. Concepts are to be regarded as the cement and bridge between work procedures and notation.

When there is a close link between work procedures, notation, and concepts, it is referred to as a method component. The concept of method component is similar to the concept of method chunk [2, 3] and the notion of method fragment [4]. A method component therefore gives instructions of how to perform a certain work step, e.g. a method component is executed through the work procedures – notation rules – and the concepts in focus. A so called "method" is often a compound of several method components into what is often referred to as a methodology [5] or a framework [1]. A framework gives a phase structure of method components, which guides us in us in terms of what to do, in what order, and what results to produce but instructions about how to do things is found in different method components.

2.2   Business Value of IT and Balanced Scorecard

During the last decades, numerous research activities from business administration, economics and computer science have addressed how to measure the business value of IT. Four typical examples of this are:

- Process-oriented approaches, like IT Business Value Metrics [20]. In process-oriented approaches the BVIT is demonstrated through process improvements. These approaches investigate how value is added to the business.
- Perceived value approaches, like IS Success Model [22]. These approaches bases BVIT evaluations on user perceptions rather than on financial indicators or measurements within technical systems.
- Project-focused approaches, like Information Economics [21]. This kind of approach basically tries to support decision making, whether an IT-project should be started by calculating a score for project alternatives.
- Scorecard-based approaches, like the balanced scorecard [13]. These approached try to include different perspective when evaluating the business value, including ,e.g., financial, process-oriented and learning perspectives.

As stated in section 1, the focus of our work is on improving the contribution of a method to the business objectives of an organization. All four types of BVIT approaches could potentially be tailored for this purpose. There are though differences between these approaches with respect to their suitability:

- the required method improvement approach has to include business value and coherence with business drivers like reduced lifecycle time or increased flexibility. These business drivers are measurable criteria reflected in control systems of many companies. Perceived value approaches do not cover these aspects sufficiently.
- method improvement requires monitoring of relevant indicators during a longer period of method use, i.e. capturing of performance indicators only once would not be sufficient. This requirement is hard to meet with project-centric approaches.
- process-oriented approaches are by nature quite specific for the individual company, as they require an understanding business processes, potential business impact and potential IT impact before starting the actual analysis of BVIT. This makes the approaches quite expensive for method improvement in terms of efforts to be invested, as methods are expected to be used in many different organizations.

Among the scorecard-based approaches, the BSC proposed by Kaplan and Norton [13] is the most established. The BSC is a *management system*, i.e. it includes measurement approaches to continuously improve performance and results.

2.3   Method Engineering

Method engineering (ME) is an expensive and knowledge intensive process, usually involving many stakeholders and results in various engineering iterations. ME is defined by [6] as the engineering discipline to design, construct and adapt methods, techniques and tools for systems development, which is also is in line with the IEEE definition of software engineering (SE) and the ISO/IEC 24744:2014 standard. In the approaches for ME and there is a need for approaches in which SE can improve the success rate [6]. A ME approach that has received a lot of interest is situational method engineering (SME) (ibid). In this paper we have applied a phase-based ME process similar to the ISO/IEC 24744:2014 standard. In this we have especially acknowledged the iterative interplay between method generation and method validation (enactment) according to Figure 1. In the ISO/IEC standard the main activities are Generation and Enactment. Generation is the act of defining and describing the method based on a defined foundation, often a meta model. Enactment is the act of validating the method through application. This standard also depicts roles like the Method engineer who is the person(s) who design, builds, extends, and maintains the method and the Developer who is the person(s) who during enactment applies the method. These two activities and roles are interlinked so that they in an interactive way can participate during both generation and enactment.
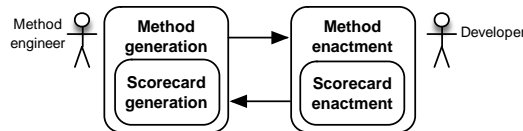

Figure 1: Method engineering process, based on ISO/IEC 24744:2014

In our ME process for the IDA-method, we developed and used a BSC as a tool for method improvement. This was done through generating and measuring different performance indicators in a BSC as part of the validation (enactment) of the method. In this the BSC has also gone through generation and validation as part of the total ME process (see figure 1), i.e., the generation and enactment of the BSC has been both interwoven in the ME-process and a parallel activity. This iterative interplay between generation and enactment for both the method and the BSC has called for a structured way to deal with this from the dimensions of both theoretical and empirical input and feedback. We followed the approach of Goldkuhl [18] proposing three levels to address during method generation and enactment (internal, theoretical, and empirical), see Table 1. This approach is similar to approaches, which also advocate theoretical and empirical dimensions of ME (c.f. [10, 19]). During our ME process we have performed different generation- and enactment activities illustrated in Table 1.

Table 1 : Generation and enactment of the IDA-method and the scorecard

	Generation	Enactment (Validation)
Internally	Reflective discussions between method engineers and developers where the emerging internal structure and content of the method and the scorecard was questioned and developed.	Evaluation of method and scorecard consistency in use in terms of structure and interrelationships of its various parts.
Theoretically	Use of a method notion to provide a conceptual structure for the method. Relating concept definitions to existing methods and established knowledge (e.g. BSC, BVIT, IDA etc.).	Comparison of the generated method against existing method notions and method theories. Analysis of method and BSC in comparison to existing method- and BVIT practices.
Empirically	Interview-based investigation for deriving method focus, conceptual foundation, and requirements. Practical application of the evolving method and BSC together with industrial partners. Development of support documentation (method- and BSC handbook) together with industrial partners.	Several test cases for evaluating the usefulness of the method and BSC in industrial cases. Industrial method use and evaluation by external parties by means of a method evaluation framework, NIMSAD. Industrial use and evaluation of method- and BSC handbook.

According to our knowledge research about method validation through the use of BSC to measure ME success for method validation is scarce. Some work can be found in relation to the actual design or construction of methods e.g. in [6]. Harmsen [16] presents a more elaborated and promising approach for using performance indicators (PI) for measuring success in IS engineering. These indicators are divided into three groups: process-related PI, product related PI, and result related PI. Even though that this research has a focus on IS engineering we believe that the same principles can be useful for method validation during ME.

3 Research approach

The research approach for development of the BSC and the engineering of the IDA-method combines design science (DS) [8] and action research (AR) [14]. In this combination, we have taken a stance in Technical Action Research (TAR) according to Wieringa and Morali [7]. In TAR the engineering process and the artifact design is the starting point and where the artifact is supposed to be validated in practice in a scaled-up sequence from test in controlled environment to full-fledged applications to solve a real practical problem in an enterprise (ibid). Our artifacts in this case are two folded, 1) the IDA-method, which is developed as a "treatment" to improve or solve information challenges in enterprises in dimensions of information supply, provision, demand, logistics etc. and 2) the BSC, which is developed as a "treatment" for method improvement. There are also earlier promising initiatives to combine DS and AR and one example of this is Action Design Research (ADR) by Sein et al. [15]. Even though that the artifact is in focus in ADR the approach still has a problem driven approach [15]. In our case the TAR approach has been more convenient since the method is developed from the notion that we have to handle information

challenges in enterprises and the BSC from the notion that there is a need for method improvement.

The work presented in this paper originates from a TAR research project with two academic and five industrial partners aiming at development of a method for IDA. The purpose with the IDA-method was to support identification, modeling and analyzing information demand as a base for development of technical and organizational solutions that provides a demand driven information provision. The ME process for the IDA-method and the BSC is described in Figure 2 below.
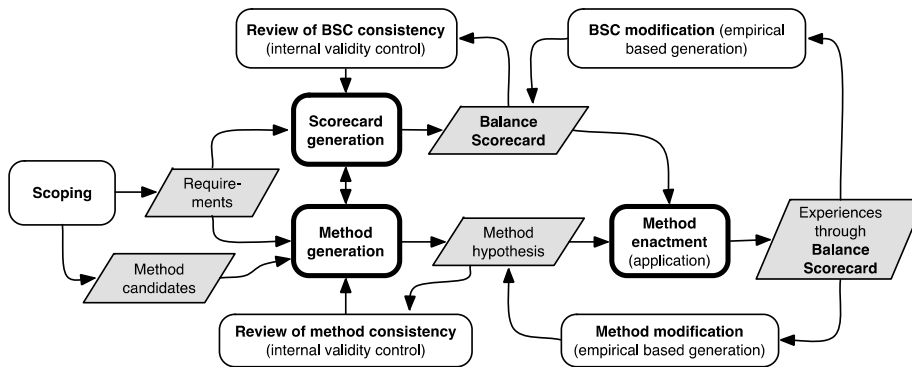

Figure 2: Generation and validation via balance scorecard of the IDA method

In this study we focus on the enactment phase and the use of BSC as a tool for method improvement. Five different IDA-cases were included:

- A metal finishing company (coordination of quality, technology and production)
- A municipality (handling of errands)
- The association for Swedish SAP users (coordination of information flow)
- A timber company (identification of information demand for test strategy)
- A gardening retailer (information demand for different organizational roles)

4   Development of BSC for Method Improvement

The BSC development is illustrated using the IDA-method which has its focus on capturing, modelling and analysing the information demand of organizational roles in order to improve the supply of information. This method has similarities to enterprise modelling methods, but it only focuses on the information demand perspective (cf. [15]). Part of the method is the use of information demand patterns, i.e. if the analysis process discovers that a certain organizational role is similar to what has been found in earlier analysis projects in other organizations, the known pattern for this role might be used and adapted. The motivation behind the narrow focus on information demand and the pattern use was to contribute to reduction of time and efforts in projects aiming at improving information flow.

Among the users of the IDA-method are consultancy companies who perform many projects aiming at improved information flows in small and medium-sized enterprises. They consider the method as a kind of resource in their "production"

process. These companies are interested to have control on the use of the method from an economic perspective, to find improvement potentials and to get at least an idea of the value for their business.

This means that the primary perspective for method improvement taken in our work is that of an organisation using the method for business purposes and aiming at improving the contribution to business objectives. Such an organisation could, for example, be a consultancy company offering analysis and optimisation services for their clients based on the method, or enterprises using the method internally for detecting and implementing change needs. In an organizational context, improvement processes usually are guided by defined goals and instruments to supervise goal achievement. The approach proposed in our work is to apply the principles of a balanced scorecard for creating a management instrument for method improvement, i.e. we will not use the original BSC perspectives and content proposed by Kaplan & Norton [15], but the process of developing such a balanced scorecard for method improvement and the general structure of goals, sub-goals, indicators, etc..

4.1   Scorecard Development Process

The main instrument for developing the scorecard was a workshop with all organizations planning to use the IDA-method. The workshop produced an initial scorecard version, which formed the basis for refinements and further development during the project. During the scorecard workshops, the following steps were taken:

The first step was to evaluate, whether the *perspectives* proposed by the original BSC approach (i.e. financial, internal business process, learning and growth, customer) are valid and appropriate for method improvement or should be changed. A starting point for identifying relevant perspectives were the strategic aims of the participating organizations. The result of this step was an initial agreement on perspectives to consider in the "method scorecard". For each perspective, *strategic goals* had to be defined and preferably quantified, as quantifying them helps to reduce the vagueness in strategic goals. Identifying strategic goals was again based on the organizations' strategy. The defined strategic goals were in a next step broken down in *sub-goals*. The objective was to define not more than 5 – 7 sub-goals per goal.

The last step related to strategic aspects was the identification of *cause-effect-relationships*. There might be strategic goals which cannot be achieved at the same time because they have conflicting elements. It is important to understand these conflicts or cause-effect relations between goals. After having covered the strategic aspects, focus was shifted on measurement issues:

For each sub-goal defined in the different perspectives, a way had to be found how to measure the current situation. For this purpose, *indicators* had to be defined contributing to capture the status with respect to the sub-goal. When defining indicators, one had to have in mind that there must be a practical way to capture these indicators. In this context, existing controlling systems or indicators (e.g. from quality management) were inspected and checked for possibilities to reuse information. For each indicator identified, the *measurement* or recording procedure was defined. A measurement procedure typically includes the way of measuring an indicator, the point in time and interval for measuring, the responsible role or person performing the measurement, how to document the measured results.

4.2   Method Validation Scorecard

The development process described in section 4.1 resulted in four different perspectives in the scorecard with the following strategic goals:

1. Method Documentation Quality: the quality of the IDA-method handbook and aids
   Goal: To have a method which is easy to train and communicate
2. Pattern Quality: the quality of the information demand (ID) patterns
   Goal: To achieve patterns of high quality applicable with the method
3. Resource Efficiency: the efficiency of the process for understanding information flow problems in enterprises and developing an appropriate solution proposal.
   Goal: More efficient resource use for the analysis including a proposal for solution
4. Solution Efficiency: the efficiency of the solution implemented in an enterprise based on using the IDA method and ID patterns
   Goal: To propose a relevant and actable solution for the case at hand

Indicator examples

Our view that a method is a guide for actions, which often are artifact-mediated actions (see section 2.1), affected the selection and definition of indicators. For brevity reasons, we will discuss sub-goals, criteria and indicators only on the basis of one of the perspectives, the method documentation quality. The overall goal "To have a method which is easy to train and communicate" was divided into several sub-goals:

- Easy to teach the method and train future modelers (transferability)
- Provide a good documentation
- Method shall support the effective development of new patterns
- Method shall take into account that patterns need continuous improvement

The criteria and indicators derived from the sub-goals are captured in a tabular way including the following information:

- What to measure, i.e. the criteria to capture. Criteria are grouped into aspects.
- Motivation of this criteria and comments (not included in table 2)
- Indicators reflecting the criteria. This is the actual value to measure.
- Indicator description: explanation related to the indicator name
- Practical implementation of capturing the indicators, i.e. how to measure, who will be responsible for measuring, when to measure and how to document the findings

Table 2 shows an excerpt of the criteria and indicator table for "method documentation quality". This excerpt is focused on "documentation quality". Further aspects in this perspective are method documentation maturity, method support for pattern use and method support for pattern extension.

The other three perspectives included the following aspects:
- Pattern Quality: applicability, technical quality, extensibility
- Resource Efficiency: analysis process, analysis result (solution), delivery
- Solution Efficiency: strategic benefits, automation benefits, transformation benefits.

Table 2: Excerpt from criteria and indicators for "method documentation quality" perspective

Aspect	Criterion	Indicator Name	Indicator Description	How to capture indicators?
Docu-mentation Quality	Learning Time	Average learning time for new analyst until "productivity"	How much time does it on average take until a person can be considered "productive"?	Captured during training sessions by method specialist
		Average learning time for new trainers until "productivity"	How much time does it on average take until a new trainer for the method can be considered "productive"?	
		Average learning time for participants in analysis projects until being able to participate in analysis projects	How much time does it on average take until a participant in analysis projects understands what she/he is contributing to?	
	Perceived quality of Method Documen-tation	Perceived quality of completeness, correctness, understandability, etc. on a suitable scale (e.g. 5 point Likert scale)		
	Documen-tation maturity	Maturity level according to review status of different stakeholder groups, like consultant, method specialist or modelling facilitator	What maturity level between "draft" and "fully validated" is assigned in method reviews by domain experts?	
		Number of improvements proposed by researchers	How many change requests were submitted for the method by researchers?	
		Number of shortcomings detected in use	How many change requests were submitted by practitioners?	

5   Method Scorecard in Use

The method scorecard was applied in two different contexts: for improving the IDA method and the enterprise modeling method 4EM [15]. When applying the scorecard in the context of IDA-method, two groups of method users have to be distinguished:

- Members of the method development team. This group obviously consisted of experts in IDA and focused on finding method improvement potential,
- Method users from outside the development team who got a training in IDA and used the method on their own shortly after the training. This group is expected to have a more independent perspective on the utility of the IDA method.

Data collected for these different groups are discussed in sections 5.1 and 5.2. In order to investigate whether the method scorecard would also be suitable for other methods than IDA, the scorecard was applied in a few 4EM cases (see section 5.3).

5.1   Scorecard use by IDA-method developers

In total 4 different members of the method development team used the scorecard during 5 different IDA cases in a time frame of 10 months. The cases addressed information flow problems in a municipality and in enterprises from retail, automotive supplier, wood-related industry, and IT industry. In each case, several modeling activities were performed, scorecard data collected and observations noted down. The observations were discussed with the other members of the method development team. As a result of the observations when using the scorecard, several adjustments were made, all of them in the first 6 months of the scorecard use:

- initially, data capturing in the cases happened based on a printed version of the document describing the scorecard. Since entering this hand-written data into a spreadsheet was tedious, a software tool was developed for data capturing. This tool offered the possibility to capture experiences and remarks in free text form.
- the solution efficiency perspective of the scorecard proved very difficult to implement and in practice not applicable. The main obstacle was that data about resource consumption, time needed for certain activities or quality of activities "before" implementing improvements detected during use of the IDA method either did not exist or were not made available due to confidentiality reasons. As a consequence, the indicators of the solution efficiency perspective were no longer captured. Instead, two new indicators were introduced: "perceived solution quality from customer perspective" and "perceived solution quality from method expert perspective". Both were captured on a 5-point scale.
- many indicators needed refinements or adjustments. An example is "average learning time for new analyst until productivity" where clarification was required whether self-study time also should be included in learning time, and whether "productivity" means being able to contribute to IDA-method use or being able to use the method self-reliantly.

The indicator data collected with the scorecard were not only evaluated during infoFLOW, but also during use of the IDA-method in later years (see also section

5.2). In every IDA use case, there were potentially four types of activities which correspond to the phases of the IDA-method: scoping, ID-context analysis, demand modeling, consolidation. Every activity type potentially requires multiple steps (i.e., activities). For each activity, scorecard indicators were captured. Example: if demand modeling required several modeling sessions with different focus areas and participants, for each of the workshops indicators were captured as separate activity.

For the presentation in this paper, we selected four indicators originating from the method documentation quality and resource efficiency perspectives of the scorecard. These four indicators were the ones preferred by the industrial partners in the infoFLOW project who intended to use the method for commercial purposes: perceived productivity, perceived method value, perceived result quality (method user) and perceived result quality (client). All indicators used the same scale: 5 – very good, 4 - good, 3- acceptable, 2 – improvements needed, 1 – poor, 0 – don't know. When preparing the data for presentation, we used two approaches:

- For all activity types in a case, we calculated the activity average for the case. Using these activity averages for a case, we calculated the overall average for a case. The case averages are shown in figure 3. The purpose of the chart was to visualize the general tendency of the method perception, here expressed in the four indicators, in order to check whether improvements made in the method handbook or the training material had any visible effect.
- The activity type averages per case are shown in figure 4. Here the intention was to see differences between activity types: where should improvements have priority?
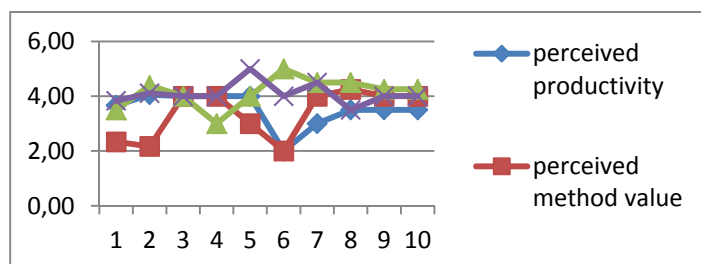

Figure 3: Case averages of selected indicators.

Figure 3 and figure 4 are based on the same cases. Cases 1 to 5 were performed by method developers; cases 6 to 10 were performed by other method users. After case 2 and case 6, a new handbook version was released. Figure 3 and 4 are meant to illustrate the indicator use in infoFLOW-2. They are not meant to prove any statistically significant developments or correlations.

The indicator development shows improvements for "perceived method value" after case 2 and case 6 when new handbook versions were released. "Perceived productivity" seems to be correlated to "perceived method value" seem to be correlated, which is not surprising. When the method was used its developers (case 1 to 5) the perceived result quality of the client was higher than of method users. When the method was used later by other method users, this is the opposite. This indicates that method developers are more critical to the results or have higher expectations.
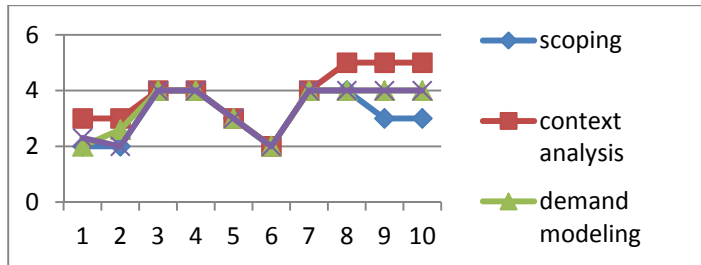
Figure 4: Activity type averages.

One of the main intentions with the activity type averages was to detect which phase should have priority when working on improvements. In cases 1 and 2, the scoping, demand modeling and consolidation needed improvement. With the new handbook published after 2, many of the problems were addressed. In demand modeling, to take one example, a notation for the demand model was included which earlier was missing. Case 5 and 6 represent the phase of transferring the method knowledge from method developer to method user. Case 5 was done in cooperation between developer and user; case 6 completely by a method user. The experiences from these first "external" uses resulted in the improvement of the handbook, i.e. from case 7 the new version was applied, which also is reflected in improved activity type averages. Currently, scoping seems to be in most need of improvement.

Table 3: Learning time for the method users

Method user (case #) Training part	Student 1 (7)	Student 2 & 3 (6)	Student 4 (8)	Student 5 (9)	Student 6 (10)
Lecture/Presentation	2	2	4	6	6
Self-study of Handbook	4	2	4	8	8
Exercise/Example	0	0	2	4	4
Coaching on case	2	0	4	4	4
Total	**8**	**4**	**14**	**22**	**22**

5.2   Scorecard use by project-external IDA-method users

In total 6 different persons were trained in the IDA-method and also used the scorecard in their information demand analysis cases, which came from logistics, manufacturing, higher education and IT industry. The scorecard indicators regarding learning time and perceived quality of the documentation were captured after the training. The other indicators were captured in every activity in each case (same as in section 5.1). 5 different cases were the basis for this paper. Section 5.1 already presented the case averages and activity type averages. Regarding learning time, table 3 shows the time invested in training the different method user, separated into lecture-like training, self-study, working on examples or coaching in real cases. The table makes clear that training was intensified for later cases which probably improved the understanding for the IDA method. This might explain the improved indicator value when comparing, e.g., case 6 and 10.

5.3 Scorecard use by 4EM-method users

Three persons used the scorecard in enterprise modeling cases with the 4EM method. The main intention was to investigate what parts of the scorecard can be used without any changes for 4EM and where adaptations need to be made. Not in scope was the comparison of IDA and 4EM based on the scorecard values.

Before the method scorecard could be used for 4EM, all perspectives, aspects and indicators were checked for suitability for 4EM:

- Method documentation quality perspective: the aspects documentation quality and method maturity could remained unchanged. Method support for pattern use and method support for pattern extension are not suitable and were removed, since 4EM does not include the use of patterns.
- Pattern quality perspective was not used – 4EM does not use patterns
- Resource efficiency perspective: all three aspects analysis process, analysis result (solution) and delivery were kept. As "analysis process" uses criteria and indicators which capture effort and duration for the different IDA method phases, these criteria had to be adapted to the activities of 4EM modeling,
- Solution efficiency perspective was not used because of the experiences made in IDA-method improvement (see 5.1)

All three 4EM modelers managed to collect data about method documentation quality and resource efficiency which confirms the feasibility of using the method scorecard for 4EM. However, in future work it should be investigated whether additional scorecard perspectives tailored to 4EM should be included. An example could be a perspective directed to participative modeling, an essential feature of 4EM.

6 Summary and Future Work

Based on the industrial project infoFLOW-2, which aimed at improving information flow in organizations, the paper presented the development process of a scorecard intended to support method improvement. The paper also presented the perspectives, aspects and (excerpts of) criteria of the method scorecard and illustrated its use for the IDA-method, and its transfer to the 4EM method. Among the conclusions to be drawn from this work are two rather "obvious" ones:

- Feasibility of scorecard development and use as support for method improvement was demonstrated. Scorecard development helped to identify what criteria and indicators were important from the organizational method users perspective.
- The transfer of the scorecard from IDA to 4EM indicates that many aspects and criteria are transferable between methods, although criteria reflecting the method phases needed adaptation. More cases are needed to confirm and refine this.

The more "hidden" conclusions are related to the utility of a scorecard: What are the actual benefits of using the scorecard? Could we have reached the same effects without the scorecard (i.e., without collecting and evaluating data)? Our impression is that the answer to these questions depends on the number of method users and cases of method use. For a method used by many persons in many cases, i.e. a sufficiently

big "sample", the data collected will help to identify elements of a method that might be candidates for improvement efforts. However, the scorecard indicators should not be considered as the "only source of truth", i.e. the scorecard should be taken as complementary means besides experience reports from method users. Section 5.1 shows an example: the indicators point at the scoping phase as a candidate for improving the method. This should be a motivation to investigate the scoping phase, but it does not mean that this part of the method really is the cause for the indicator values – there might be other causes, like e.g., the qualification of the modelers for "scoping" or the measurement procedure for the indicators might be inadequate.

Furthermore, some criteria and indicators of the scorecard need further investigation regarding their usefulness. Example is the average time required for the different phases of the IDA method. This time is partly dependent on modeler and complexity of the case. But if there are many projects and different modelers, the development tendency of the average values of this indicator can be relevant.

Our preliminary recommendations regarding the method scorecard can be summarized as follows:

- use the scorecard only for methods with many users and cases
- for indicators addressing the time or effort required for certain activities: find way to normalize the complexity of these different activities
- consider to reduce number of indicators, e.g., to 5 per perspective
- Use tool support for capturing and evaluating indicators
- Use scorecard as complementary means for method evaluation and improvement only. Very valuable information for improvement of methods usually comes from the method users
- Indicators can help in method evolution management

Future work will on the one hand consist of continued data collection regarding the IDA method, which will probably lead to further development of the scorecard, and further investigation of transferability of the scorecard to other methods. Furthermore, more work is needed on understanding from what number of method users and cases a scorecard use is recommendable. It also has to be investigated if a scorecard designed for method improvement in organisational purposes also can be applied as instrument in method engineering. This is to a large extent a question of generalisability of scorecard perspectives and indicators, i.e. is the scorecard for a specific organisational context also (in total or parts) valid for the general use of the method?

References

1. U. Seigerroth: Enterprise Modeling and Enterprise Architecture: the constituents of transformation and alignment of Business and IT, International journal of IT/Business Alignment and Governance (IJITBAG), ISSN 1947-9611, vol. 2, 2011.
2. J. Ralyté, P. Backlund, H. Kühn, M. A. Jeusfeld: Method Chunks for Interoperability, in Conceptual Modeling - ER 2006, vol. 4215, LNCS, 2006, pp. 339-353. Available: http://dx.doi.org/10.1007/11901181_26
3. I. Mirbel, J. Ralyté: Situational method engineering: combining assembly-based and roadmap-driven approaches, Requirements Engineering, vol. 11(1), pp. 58–78, 2006. Available: http://dx.doi.org/10.1007/s00766-005-0019-0

4. S. Brinkkemper: Method engineering: engineering of information systems development methods and tools, Information and Software Technology, vol. 38(4), pp. 275-280, 1996. Available: http://dx.doi.org/10.1016/0950-5849(95)01059-9

5. D. E. Avison, G. Fitzgerald: Information Systems Development: Methodologies, Techniques, and Tools, Maidenhead, UK: McGraw-Hill Education, 1995

6. B. Henderson-Sellers, J. Ralyté, P. Ågerfalk, M. Rossi: Situational Method Engineering: Springer Verlag Berlin Heidelberg (2014)

7. R. Wieringa, Morali A. (2012) Technical Action Research as a Validation Method in Information Systems Design Science, LNCS 7286, Springer-Verlag Berlin Heidelberg 2012

8. Hevner A. R., March S. T., Park J., Ram S. (2004). Design Science in Information Systems Research. *MIS Quarterly, 28*(1), 75-105.

9. March S.T., Smith G. (1995) Design and Natural Science Research on Information Technologies, *Decision Support Systems*, 15(4), pp 251-266.

10. Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry (Vol. 75). Sage.

11. Lind M., Seigerroth U., Forsgren O., Hjalmarsson A. (2008) Co-design as social constructive pragmatism, AIS Special Interest Group on Pragmatist IS Research (SIGPrag 2008) at International Conference on Information Systems (ICIS2008), France.

12. Lundqvist, M.; Sandkuhl, K.; Seigerroth, U. (2011) Modelling Information Demand in an Enterprise Context: Method, Notation and Lessons Learned. International Journal Systems Modeling and Design, Vol. 2 (3), IGI Publishing, pp. 74-96, 2011.

13. Kaplan, R. S., & Norton, D. P. (1996). The balanced scorecard: translating strategy into action. Harvard Business Press.

14. Susman, G. I., & Evered, R. D. (1978). An Assessment of the Scientific Merits of Action Research. *Administrative Science Quarterly, 23*(4), 582-603.

15. Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action Design Research. *MIS Quarterly,,35*(1), 37-56.2).

16. Harmsen A. F. (1997) Situational Method Engineering, Doctoral dissertation University of Twente, ISBN: 90-75498-10-1, Moret Ernst & Young, Utrecht, The Netherlands.

17. Sandkuhl, K.; Stirna, J.; Persson, A.; Wißotzki, M. (2014) Enterprise Modeling: Tackling Business Challenges with the 4EM Method. Springer. ISBN 978-3662437247.

18. Goldkuhl, G. (1999) The Grounding of Usable Knowledge: An Inquiry in the Epistemology of Action Knowledge. In CMTO Research Papers, No. 1999:03, Linköping University.

19. Siau, K. & Rossi, M. (2008) Evaluating techniques for system analysis and design modelling methods – a review and comparative analysis. In Information System Journal. Blackwell Publishing Ltd (2008)

20. Mooney J., Gurbaxani V., Kraemer K. (1995). A Process Oriented Framework for Assessing the business value of information technology. In: Proceedings of the 16th International Conference on Information Systems, Amsterdam, pp. 17-27.

21. Parker M., Benson R. (1998). Information Economics, Prentice-Hall, Englewood Cliffs, NJ.

22. DeLone W., McLean E. (1992). Information system success: the quest for the dependent variable. Information Systems Research, Vol. 3, No. 1, pp. 60-95.