University of Augsburg
Prof. Dr. Hans Ulrich Buhl
Research Center
Finance & Information Management
Department of Information Systems
Engineering & Financial Management

Discussion Paper WI-253

How to Select Measures for Decision Support Systems – An Optimization Approach Integrating Informational and Economic Objectives

by

Max Röglinger

in: S. Newell, E. Whitley, N. Pouloudi, J. Wareham, L. Mathiassen, eds.,
Proceedings of the 17th European Conference on Information Systems,
ECIS, Verona, Juni 2009, p. 2856-2868

Universität Augsburg, D-86135 Augsburg
Visitors: Universitätsstr. 12, 86159 Augsburg
Phone: +49 821 598-4801 (Fax: -4899)
www.fim-online.eu

HOW TO SELECT MEASURES FOR DECISION SUPPORT SYSTEMS – AN OPTIMIZATION APPROACH INTEGRATING INFORMATIONAL AND ECONOMIC OBJECTIVES

The author(s)

Abstract

It is still an open issue of designing and adapting (data-driven) decision support systems and data warehouses to determine relevant content and in particular (performance) measures. In fact, some classic approaches to information requirements determination, such as Rockart's critical success factors method, provide valuable assistance with structuring decision makers' information requirements and identifying thematically appropriate measures. In many cases, however, it remains unclear which and how many measures should eventually be used. Therefore, an optimization model is presented that integrates informational and economic objectives. The model incorporates (statistic) interdependencies among measures – i. e. the information they provide about one another –, decision makers' and reporting tools' ability of coping with information complexity as well as negative economic effects due to measure selection and usage. We show that in general the selection policies of all-or-none or the-more-the-better are not reasonable although they are often conducted in business practice. Finally, the model's application is illustrated by the German business-to-business sales organization of a global electronics and electrical engineering company as example.

Keywords: decision support systems, information overload, critical success factors, design research.

1 MOTIVATION AND OBJECT OF RESEARCH

Due to the complexity of intra- and extraorganizational structures, it is impossible for decision makers in general – and executives in particular – to continuously monitor all fields of action that possibly require intervention. With reports containing in average 12,000 to 15,000 data points based on (performance) measures, i. e. key figures or indicators, information proliferation makes it even harder to focus on decision-relevant information (Axson 2007). Some measures more or less significantly influence the complexity of reports and the amount of time needed to understand them. Hence, a central problem in the design and adaptation of (data-driven) decision support systems (DSS) and data warehouses (Alter 1980, Inmon 2005, Power 2002) still is to determine relevant fields of action and to select appropriate measures (Eccles 1991, NIST 2008, Watson et al. 1993). Particularly the latter requires formal research (Evans 2004).

Some classic approaches to information requirements determination (IRD), such as Rockart's critical success factors (CSF) method (1979), provide valuable assistance with structuring decision makers' information requirements (IR) and identifying thematically appropriate measures. However, these measures are often too many and it is unclear which should eventually be used. In this respect, decision makers' cognitive restrictions (Browne et al. 2002, Davis 1982), limitations of reporting tools, such as management cockpits and dashboards (Sisfontes-Monge 2007), as well as negative monetary implications need to be considered. As for measure selection in particular, there are additional deficiencies with respect to whether the selection process is intersubjectively comprehensible, decision makers can participate systematically, and (statistic) interdependencies among measures (e. g. quantifiable by means of correlation or contingency coefficients) are considered. In business practice, these deficiencies can result in that measure selection is based on "gut instinct", that many time-consuming interviews are conducted, and that the utility of selected measures remains doubtful.

Therefore, the research question is: *Which and how many measures shall be selected from a preselected set of thematically appropriate measures in order to provide decision makers with optimal information as regards informational and economic objectives?*

The paper relies on a design-oriented, formal, and deductive approach (Hevner et al. 2004). Section 2 compares existing approaches with respect to general requirements and identifies the research gap. Section 3 proposes an optimization model as artifact. Section 4 evaluates the optimization model by illustrating its application in business practice and by assessing how it meets the general requirements outlined above. Section 5 summarizes the results and points out future research.

2 RELATED WORK

Currently, measures are often contained in performance measurement systems (PMS). In management accounting and operations management literature, there is a range of requirements on PMS (e. g. Artley et al. 2001, Caplice et al. 1995, Neely et al. 1995). Accordingly, PMS are expected to capture all relevant constituencies of a specific field of action (*completeness, R.1*), to encompass a manageable amount of measures (*clarity, R.2*), and to transfer the overall business strategy to decision makers (*vertical integration, R.3*). The process of measure selection should be intersubjectively comprehensible (*intersubjectivity, R.4*), consider (statistic) interdependencies among measures (*interdependencies, R.5*), and involve domain experts (*participation, R.6*). Some requirements are complementary, while others are competing. For instance, completeness and clarity are competing, whereas interdependencies and clarity are complementary. The requirements are also somehow vague due to prosaic formulation. Nevertheless, in the author's opinion they provide basic assistance with comparing existing approaches and with identifying the research gap (see Table 1 where completeness is omitted as it is not addressed by any approach).

	Clarity (*R.2*)	Vertical integration (*R.3*)	Intersubjectivity (*R.4*)	Interdependencies (*R.5*)	Participation (*R.6*)
Giorgini et al. (2008)	No maximum of figures	By goal analysis	Subjective mapping of measures to goals	Isolated consideration	-
Neely et al. (2000)	No maximum of measures	By business strategy	Partial intersubjectivity via checklists	Postulated, but not elaborated	-
Liebetruth et al. (2006)	Arbitrary maximum of measures	By CSF	Partial intersubjectivity via optimization model	Postulated, but not elaborated	Assignment of utility scores
Rockart (1979)	No maximum of measures	By CSF	Subjective mapping of measures to CSF	Isolated consideration	Explorative interviews
Axson (2007)	No maximum of measures	By CSF and business strategy	Subjective mapping of measures to CSF	Isolated consideration	Explorative interviews and "games"

Table 1. Comparison of existing approaches to measure selection

Giorgini et al. (2008) present a goal-oriented approach to determine IR for data warehouses that considers the organizational environment and decision makers' needs. Neely et al. (2000) advocate a selection of measures in terms of a cost-benefit-analysis. Liebetruth et al. (2006) present a linear optimization model with which a utility-optimal subset of measures can be chosen from a set of preselected and thematically appropriate measures. Rockart (1979) shows how decision makers' IR can be structured and reduced to a few essential fields of action, the so-called CSF, each of which is monitored by several measures (see also Leidecker et al. 1984). Axson (2007) extends CSF analysis by incorporating additional interactive elements, distinguishing primary and supporting measures as well as vaguely postulating "minimal confusion".

The following findings are noteworthy: Almost all approaches neglect clarity (*R.2*) as they do not specify how many measures are to be selected. One approach allows to set a maximum number of measures. This is arbitrary and considers neither the decision makers' information processing capacity nor economic implications. All approaches are vertically integrated by linking measures with CSF, business strategy, or goals (*R.3*). Moreover, measure selection is (at least) partially subjective (*R.4*). Interdependencies among measures are not considered (*R.5*). Most approaches involve decision makers by means of explorative elements (e. g. interviews or games) (*R.6*). Summing up, there is a primary research gap with respect to clarity (*R.2*) and interdependencies (*R.5*). Furthermore, there still is potential for improvement with respect to intersubjectivity (*R.4*) and participation (*R.6*).

We focus on the primary research gap. In the end, this will also ameliorate the other requirements. We adopt the ideas of optimization and preselection of thematically appropriate measures from Liebetruth et al. (2006), the structuring momentum of CSF from Rockart (1979), and the idea of explicitly incorporating negative economic implications from Neely et al. (2000). Our contribution is that we formally address the trade-off between provided information, information complexity, and negative economic implications to determine which and how many measures should be selected optimally.

3 AN OPTIMIZATION MODEL FOR MEASURE SELECTION

Consider a company where the reporting has historically grown and multiple (data-driven) DSS and data warehouses are in use. In order to react on its decision makers' demand for clear information, the company launches a project for implementing a consolidated DSS. Two essential steps in this project are: (1) structuring the decision makers' IR into relevant fields of action and (2) (pre-)selecting thematically appropriate measures from the existing systems – assuming no new measures will be added. In most cases, it will not be reasonable to integrate all preselected measures – nor even only those desired by the decision makers (e. g. Ackoff 1967, Davis 1982). This is for several reasons: some measures may (partially) "overlap" due to (statistic) interdependencies, decision makers can only cope with restricted information complexity, and customizing as well as maintaining reporting tools is expensive. It is thus advisable to analyze in advance which fraction of the preselected measures the consolidated DSS should contain. Whereas above thematically appropriate measures had to be (pre-)selected, here measures are of interest that together provide much information about other measures. As indicated, two perspectives are important here: the *economic* and the *informational* perspective. While the former is indispensable when investing in IT, the latter is necessary as DSS primarily aim at supporting decision processes by supplying decision-relevant information (e. g. Power 2002).

In order to determine the optimal fraction of measures, we propose an optimization model. Though being inherently discrete, the problem of measure selection can be interpreted as approx. continuous for sufficiently many measures. This allows to determine algebraic solutions and to gain general insights. We maintain the affiliation with the original problem setting and make reasoning about functions more illustrative by also using discrete examples. A basic model for the informational perspective is proposed in section 3.1 and extended by the economic perspective in section 3.2.

3.1 A basic model for the informational perspective

Let us first consider the informational perspective, where information has no price. Selecting one of the preselected measures provides information about the measure itself – as it becomes known – and about non-selected measures – due to (statistic) interdependencies. This creates informational utility. The more strongly a measure interdepends with non-selected measures, the more informational utility it creates because it allows to estimate their values more reliably. There are also negative informational effects of selecting measures. Due to increasing information complexity, each additional measure makes it harder to cognitively process the entire amount of information. This creates informational disutility. Thus, there is an *informational trade-off*. The question is: Up to which optimal fraction of measures does the utility due to more information justify the disutility due to higher information complexity? The optimization model relies on the following assumptions:

A.1: There is a given finite set of measures, which have been preselected ex ante with respect to thematic appropriateness. Between some measures, there are meaningfully interpretable pairwise (statistic) interdependencies, that is, selected measures provide information about (the values of) non-selected measures. All measures together satisfy the decision makers' information requirements and provide complete information. Moreover, all measures together cause highest complexity.

A.2: The fraction of the preselected measures that will be integrated into the consolidated DSS, $x \in [0;1]$, is infinitely divisible (see discussion above). With $x = 0$, no measures are selected. With $x = 1$, all measures are selected.

A.3: $U_{info}(x)$ represents the informational utility due to the information that a fraction of selected measures provides about itself and non-selected measures. $D_{info}(x)$ represents the informational disutility due to information complexity. Both are functions of x and can be forecast ex ante.

On these assumptions, the informationally optimal fraction of measures $x^{opt}{}_{info}$ can be determined by optimizing the difference between $U_{info}(x)$ with $D_{info}(x)$. This difference is also called informational net utility $U_{info,net}(x)$. The corresponding objective function is given by:

$$U_{info,net}(x) = U_{info}(x) - D_{info}(x) = \max! \tag{1}$$

In order to formalize the optimization model, $U_{info}(x)$ and $D_{info}(x)$ are examined. We start with $U_{info}(x)$. If a (rational) decision maker were restricted to select only one measure, he would select the measure with the highest *individual* informational utility – say m_1 –, i. e. the measure that in sum interdepends most strongly with the non-selected measures. If the decision maker were allowed to select two measures, he would take those that create the highest *joint* informational utility – say m_2 and m_3. In general, this joint informational utility is higher than the individual utility of m_1. This is because either

m_1 is kept (as m_2 or m_3) and another measure is added or m_1 is discarded and two other measures with higher joint utility are chosen. The only exception is if all measures interdepend perfectly. In this case already one measure alone – no matter which – provides complete information. In general, the joint informational utility of m_2 and m_3 is smaller than the sum of both individual utility values. This is because interdependencies cause "informational overlap". To put it more precisely: If we only consider m_2 and m_3, the joint utility of knowing both m_2 and m_3 is smaller than the sum of the individual utility values due to knowing m_2 (or m_3) and its interdependency with m_3 (or m_2). If m_2 and m_3 interdepend, m_2 provides information about m_3 – and vice versa. The only exception is if m_2 and m_3 are (statistically) independent of each other. In this case, the joint utility equals the sum of both individual utility values. If we consider all non-selected measures, the joint interdependency-induced utility of m_2 and m_3 is smaller than the sum of the individual interdependency-induced utility values of m_2 and m_3. For each non-selected measure, the strongest interdependency will be used to estimate its value. The only exception is if m_2 and m_3 are independent of all non-selected measures or interdepend with disjoint subsets of non-selected measures. With many measures at hand, this is rather unlikely. Hence, the marginal utility of selecting m_2 and m_3 (two measures) compared to selecting m_1 (one measure) is smaller than (and exceptionally equal to) the marginal utility of selecting m_1 (one measure) compared to selecting zero measures. This holds for any number of measures. Hence, the more measures are selected – i. e. the higher x is –, the higher is the joint informational utility and the less is the marginal utility. In mathematical terms, $U_{info}(x)$ is increasing ($\partial(U_{info}(\mathrm{x}))/\partial x \leq 0$) and concave ($\partial^2(U_{info}(x))/\partial x^2 \leq 0$). If we neglect the discussed exceptions and treat $U_{info}(x)$ as *strictly* increasing and concave, it may be formalized in a simplifying manner as follows:

$$U_{info}(x) = x^{\alpha} \cdot A \text{ with } \alpha \in \,]0;1] \text{ and } A \in IR^{+} \tag{2}$$

Selecting no measures provides no information ($U_{info}(0) = 0$), whereas – according to *A.1* – selecting all measures provides complete information ($U_{info}(1) = A$). The constant A represents the decision makers' present-value monetary equivalent of complete information, that is, the amount of money they are willing to pay at the moment of measure selection for complete information during the planning horizon, i. e. as long as the selected measures are in use. Reasoning from an informational perspective, A represents the value of information by itself. It does not incorporate payments e. g. for data collection. The transformation into monetary units enables to integrate the economic perspective later. The diminishing marginal utility, which was introduced above and is caused by a higher fraction of measures, is formalized by the fact that the exponent α is restricted to]0;1]. This also excludes the case where all measures interdepend perfectly, which would lead to a non-realistic course of $U_{info}(x)$. A value of α close to 0 is appropriate if all preselected measures interdepend rather strongly. Therefore, very few measures already create almost complete information. A value of α close to 1 is appropriate if all measures are rather independent, that is, the marginal utility is rather constant. A mean value of α indicates that the measures split into several groups with strong intra-group and weak inter-group interdependencies. The higher the value of α, the more (and the smaller) groups tend to exist.

The objective function's second component represents the disutility created by information complexity $D_{info}(x)$. It intuitively holds that the more measures are selected, the more complex it is to cognitively process them. Mathematically spoken, $D_{info}(x)$ increases with x. According to cognitive sciences (e. g. Duncan 1980, Miller 1956), the amount of information becomes overproportionally more complex when the fraction of measures increases. Hence, a higher fraction x is also characterized by an increasing marginal disutility with respect to $D_{info}(x)$. In summary, $D_{info}(x)$ is strictly increasing ($\partial(D_{info}(x))/\partial x > 0$) and strictly convex ($\partial^2(D_{info}(x))/\partial x^2 > 0$). This can be formalized as follows:

$$D_{info}(x) = x^{\beta} \cdot B \text{ with } \beta \in \,]1;\infty[\text{ and } B \in IR^{+} \tag{3}$$

Selecting no measures does not lead to complexity ($D_{info}(0) = 0$), whereas – according to *A.1* – selecting all measures leads to highest complexity ($D_{info}(1) = B$). The constant B represents the decision makers' present-value monetary equivalent of understanding complete information during the planning horizon. The increasing marginal disutility, which was introduced above and is caused by a higher fraction of measures, is formalized by the fact that the exponent β is restricted to]1;∞[. Its value depends on the decision makers' and employed reporting tools' ability of coping with information complexity[1]. A value close to 1 is appropriate if the decision makers already have serious problems with processing few measures and/or the employed reporting tools are restricted to a few measures. The higher the value of β, the less decision makers are susceptible to information complexity and/or the more powerful the employed reporting tools are. Based on (1) to (3), the optimization model is as follows:

$$\begin{aligned} &\text{Maximize} && U_{info,net}(x) = U_{info}(x) - D_{info}(x) = x^{\alpha}\cdot A - x^{\beta}\cdot B \\ &\text{w. r. t.} && x \in [0;1] \end{aligned} \qquad (4)$$

A mathematical analysis shows that $U_{info,net}(x)$ strictly increases until $x^*_{info} = [(A\cdot\alpha)/(B\cdot\beta)]^{1/(\beta-\alpha)}$. Up to that fraction, each additional measure provides more additional information than it causes additional complexity. Beyond, $U_{info,net}(x)$ strictly decreases. Each additional measure then causes more additional complexity than it provides additional information. As x is restricted to [0;1], the optimal fraction is $x^{opt}_{info} = \min\{x^*_{info};1\}$. Due to the concave course of $U_{info,net}(x)$, a border solution such as $x^{opt}_{info} = 1$ only occurs on rare occasions (see below).

Two interesting questions are: How is the decision makers' attitude towards complete information and highest complexity reflected in A and B? How do both parameters ceteris paribus affect the course of $U_{info,net}(x)$ and the position of x^{opt}_{info}? The following case differentiation is also depicted in Figure 1. If $A = B$, complete information creates as much utility as highest complexity creates disutility. Decision makers then are indifferent between making decisions based on zero measures or based on all preselected measures. The optimal fraction is $x^{opt}_{info} = (\alpha/\beta)^{1/(\beta-\alpha)}$ and only depends on α and β. If $A < B$, complete information creates less utility than highest complexity creates disutility. Decision makers prefer making decisions based on zero measures to making decisions based on all measures. The optimal fraction x^{opt}_{info} is ceteris paribus smaller than in the first case. If $A > B$, complete information creates more utility than highest complexity creates disutility. Decision makers prefer making decisions based on all measures to making decisions based on zero measures. $U_{info,net}(x)$ becomes zero only once in [0;1]. The optimal fraction x^{opt}_{info} is ceteris paribus higher than in the first case. For certain constellations of α and β (see Figure 1 on the right), $U_{info,net}(x)$ could have its maximum x^*_{info} outside the interval [0;1]. With x being restricted to this interval, the optimal fraction then is $x^{opt}_{info} = 1$. This is the only case where it may be informationally optimal to select all measures.

3.2 An extended model for the informational and the economic perspective

In reality, information is not for free. Hence, it is necessary to integrate an economic perspective. In order to support decision makers, measures need to be compiled into reporting tools (e. g. management cockpits and dashboards). These need to be customized and maintained during their time in use. Abstracting from fixed costs, this leads to one-time and continuous payments. Both create economic disutility and influence measure selection. Thus, there is a *joint informational and economic trade-off*. The question is: Up to which optimal fraction of measures does the additional informational net utility justify the additional economic disutility due to higher present-value payments for customization and maintenance? The extended model additionally relies on the following assumptions:

A.4: All preselected measures are implemented and their values can be extracted automatically from the respective application systems. The consolidated DSS will be connected to the existing application systems.

[1] To simplify matters, β is viewed as average value of how well decision makers/reporting tools are able to cope with information complexity. Of course, it can be further refined with respect to different groups/types or even individual decision makers/reporting tools.

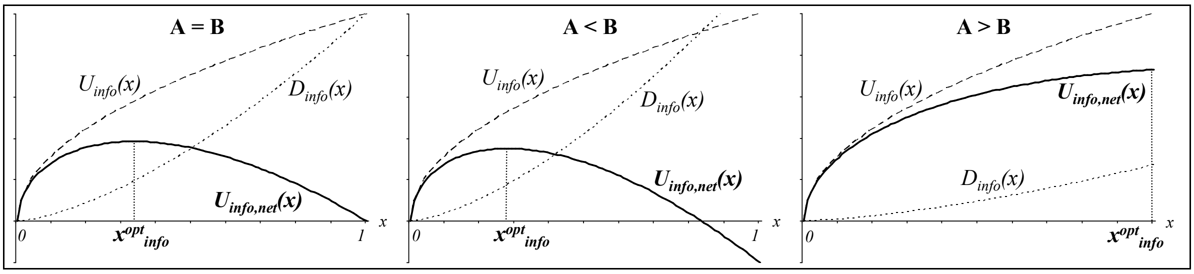

Figure 1. *Exemplary courses of $U_{info}(x)$, $D_{info}(x)$, and $U_{info,net}(x)$*

A.5: $D_{econ}(x)$ is the economic disutility due to the present-value payments for customizing and maintaining reporting tools. It is a function of x and can be forecast ex ante.

On the assumptions *A.1* to *A.5*, the informationally *and* economically optimal fraction $x^{opt}_{info+econ}$ can be determined by optimizing the difference between $U_{info,net}(x)$ and $D_{econ}(x)$. This difference is also called joint informational and economic net utility $U_{info+econ,net}(x)$. The objective function is given by:

$$U_{info+econ,net}(x) = U_{info,net}(x) - D_{econ}(x) = U_{info}(x) - D_{info}(x) - D_{econ}(x) = \max! \quad (5)$$

To formulate the extended optimization model, $D_{econ}(x)$ is examined. According to *A.4*, payments for systems integration and data collection need not be considered. The more measures are selected, the more time consuming – and expensive – it is to initially customize reporting tools. Imagine the selected measures had to be integrated into a dashboard. If only one measure is selected, the dashboard can be customized easily. If two measures are selected, an overall layout is more difficult (but still easy) to find. The more measures are selected, the overproportionally more time-consuming – and expensive – is it to find an adequate overall layout. This includes choosing among different visualization elements, adapting their size, trying different layouts, or – in the worst case – changing the reporting tool. This also applies to the present-value payments for maintaining reporting tools. These arise e. g. for updating ETL procedures, assuring data quality, or changing selected measures. According to Axson (2007), in average 15 to 20 % of the selected measures will have to be changed during the first year, 10 to 15 % in the following years. Hence, the more measures are selected – that is, the higher x is –, the higher is the economic disutility and the higher is the marginal disutility. Mathematically spoken, $D_{econ}(x)$ is strictly increasing ($\partial(D_{econ}(x))/\partial x>0$) and strictly convex ($\partial^2(D_{econ}(x))/\partial x^2>0$). This may be formalized as follows:

$$D_{econ}(x) = x^{\gamma}\cdot C \text{ with } \gamma \in]1;\infty[\text{ and } C \in IR^{+} \quad (6)$$

Selecting no measures does not lead to payments ($D_{econ}(0) = 0$), whereas selecting all measures leads to highest payments ($D_{econ}(1) = C$). The constant C represents the highest amount of present-value payments due to customization and maintenance of reporting tools. The increasing marginal disutility, which was introduced above and is caused by a higher fraction of measures, is formalized by the fact that the exponent γ is restricted to $]1;\infty[$. A value close to 1 is appropriate if a small fraction of measures already leads to high payments and each measure causes approx. the same marginal disutility. The higher γ is, the less payments and marginal disutility causes a small fraction of measures and the higher is the marginal disutility of higher fractions. Based on (4) to (6), the extended optimization model is as follows:

$$\begin{aligned} &\text{Maximize} \quad U_{info+econ,net}(x) = U_{info}(x) - D_{info}(x) - D_{econ}(x) = x^{\alpha}\cdot A - x^{\beta}\cdot B - x^{\gamma}\cdot C \\ &\text{w. r. t.} \quad x \in [0;1] \end{aligned} \quad (7)$$

Although there is no general algebraic solution, the course of $U_{info+econ,net}(x)$ and the position of $x^{opt}_{info+econ}$ can be discussed with respect to the component functions (see Figure 2). As $U_{info,net}(x)$ is concave and $D_{econ}(x)$ is convex, $U_{info+econ,net}(x)$ is concave with one global maximum at $x^{*}_{info+econ}$. As $U_{info+econ,net}(x)$ equals $U_{info,net}(x)$ diminished by $D_{econ}(x)$, the joint informational and economic optimum

$x^{opt}_{info+econ}$ is smaller than or equal to x^{opt}_{info}, that is, $x^{opt}_{info+econ} \in\ ]0;x^{opt}_{info}]$. This is reasonable because x^{opt}_{info} is determined on the assumption that information has no price. If $D_{econ}(x)$ is close to zero – e. g. for large γ and/or very small C –, measures can be selected almost negligent of negative economic effects. Then $x^{opt}_{info+econ}$ and x^{opt}_{info} are approx. equal. If $D_{info}(x)$ is close to zero – e. g. for large β and/or very small B –, decision makers are hardly susceptible to information complexity and/or powerful reporting tools are employed. Then an approx. solution is $x^{opt}_{info+econ} \approx (A\cdot\alpha/C\cdot\gamma)^{1/(\gamma-\alpha)}$. In this case, analogous to above, the relationship between the decision makers' subjective attitude towards complete information and highest (present-value) payments can be analyzed with respect to A and C.

Concluding, the optimization model allows determining the optimal fraction of measures to be chosen from a preselected set of thematically appropriate measures. Accordingly, those figures are selected that together create the highest informational utility. The model integrates an informational and an economic perspective. The former reflects the decision makers' attitude towards information and information complexity. The latter considers present-value payments for customizing and maintaining reporting tools. It could be shown that, in general, the selection policies of all-or-none or the-more-the-better, which are often implemented in business practice, are reasonable neither from an informational nor from a joint informational and economic perspective. What makes sense instead is a differentiated and balanced selection of measures.

4 EVALUATION

4.1 Applying the optimization model in business practice

The optimization model was applied in the context of a project at the German business-to-business sales organization of a global electronics and electrical engineering company. As there were only few measures, the model could be applied in a *discretized* form. If there had been very much measures, we would have had to evaluate a manageable subset in order to infer the continuous functions introduced above. Due to confidentiality, all data is anonymized and modified. Yet the principal results still hold. The project's overall goals were to better support the sales force, to reduce IT operation costs, and to modernize sales reporting. As for the first two goals, the company decided to introduce a single CRM system and to harmonize the application landscape, which consisted of more than one hundred division-specific legacy systems. Salespeople in average spent about 15 % of their time on multiple work. Large parts of the cross-selling potential – particularly across divisions – were not tapped. The reporting mainly consisted of financial and lagging measures such as volume of sales and price margin. It was to be modernized with respect to *non-monetary* and *leading* measures. Our task was to structure the salespeople's IR into CSF and to select appropriate measures. At first, candidate CSF were identified by explorative interviews with 10 sales managers and senior members of the company's CRM board. Sales managers had usually worked as sales representatives for several years and were supposed to provide valuable hints with respect to IR and sales reporting. For each candidate CSF, several items were identified and compiled into a five-point Likert scale-based questionnaire,

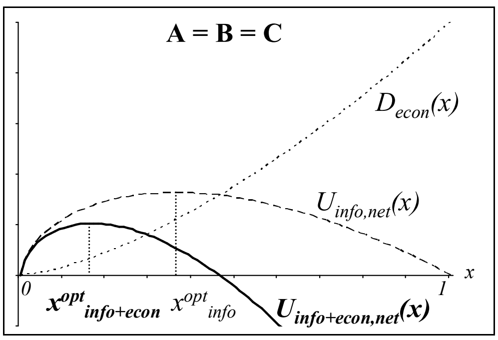

Figure 2. Exemplary course of $U_{info+econ,net}(x)$ for $A = B = C$

which was presented to 25 sales managers after a pretest. All in all, 10 CSF were identified for 3 dimensions, namely organizational structures and processes (e. g. long-term customer care, cross-divisional cooperation), salespeople's skills and knowledge (e. g. with respect to installed base and competitors' portfolios), and IT functionality (e. g. integration with office communication software, IT-based planning of sales calls).

The CSF "cross-divisional cooperation" will serve as example. Together with the sales managers, we retrieved 8 thematically appropriate measures. These were: fraction of converted leads[2] from other divisions (*%_leads*), average overall time spent on creating leads for other divisions (*∅_T_leads*), average time spent on creating one lead (*∅_T_lead*), number of trainings on other divisions' portfolios (*#_trainings*), number of meetings with colleagues from other divisions (*#_meetings*), number of sales calls with colleagues from other divisions (*#_calls*), number of shared customers (*#_customers*), number of bids for customers of other divisions (*#_bids*). All measures had existed for several years and were reported monthly on a sales manager's granularity.

First, we assessed informational utility $U_{info}(x)$. The exponent α – which indicates how the measures interdepend – was operationalized based on Pearson's correlation coefficient. That is, we treated the interdependencies as pairwise, symmetric, and linear. We accepted this simplification as the correlation coefficient is an intuitive, widely used, and relatively easy-to-compute measure. Moreover, linear interdependencies are often considered as sufficiently good approximations for many economic settings (Edwards 1976). This turned out to be useful because, due to missing hierarchical and logic structures, the existence and strength of interdependencies among non-monetary and leading measures often need to be assessed empirically e. g. by interviewing domain experts and analysing historical data (Küpper 2005). Let $M = \{m_1, m_2, ..., m_n\}$ comprise n preselected and thematically appropriate (metrically scaled) measures between some of which there are meaningfully interpretable correlations. We considered absolute values as the correlation coefficient's algebraic sign only indicates direction, not strength. The values are represented as $n{\times}n$-matrix C_M where c_{ij} indicates how strong m_i and m_j correlate and where $c_{ij} = 0$ if m_i and m_j are statistically independent or their correlation cannot be meaningfully interpreted ($0 \le i, j \le n, i \ne j$). The individual correlations of a measure m_i equal the i-th column vector of C_M.

$$c_i = (c_{i1}, c_{i2}, ..., c_{in})^T \tag{8}$$

The joint correlations of multiple selected measures $m_1,...,m_i$ (without loss of generality) are also represented as vector $c_{1,...,i}$. The elements of all selected measures are 1 (perfect autocorrelation). The element of each non-selected measure indicates the strongest correlation with any selected measure. This is reasonable because, if decision makers want to estimate the value of a non-selected measure m_j, they will use the selected measure that correlates most strongly with m_j.

$$c_{1,...,i} = (\max\{c_{11},...,c_{i1}\}, \max\{c_{12},...,c_{i2}\}, ..., \max\{c_{1n}, ...,c_{in}\})^T \tag{9}$$

The concept of joint correlations enables to formalize a discretized informational utility as function of x. We need the highest joint correlation of $x{\cdot}n$ measures. It is determined by calculating the highest scalar product value $\langle 1,c^x \rangle$ where 1 is an n-vector $(1, 1, ..., 1)^T$ and c^x is the joint correlations vector of $x{\cdot}n$ arbitrary measures. Dividing the scalar by n normalizes it to [0;1]. This operationalization can be interpreted as a monetized mean absolute correlation.

$$U_{info}(x) = [\max \{ \langle 1,c^x \rangle \mid x{\cdot}n \text{ measures are selected} \} / n] \cdot A \tag{10}$$

After CSF analysis, some sales managers were asked between which measures meaningfully interpretable correlations existed. Only their strengths were calculated based on historical data. Table 3 shows the results with the light grey cells marking excluded correlations. Concerning the

[2] In the CRM context, a lead represents a hint with low degree of maturity from inside or outside one's division that refers to a potential customer or project opportunity.

	%_leads	*Ø_T_leads*	*Ø_T_lead*	*#_trainings*	*#_meetings*	*#_calls*	*#_customers*	*#_bids*
%_leads	**1.00**	0.43	0.83	0.67	0.00	0.42	0.34	0.96
Ø_T_leads	0.43	**1.00**	0.34	0.00	0.00	0.00	0.12	0.54
Ø_T_lead	0.83	0.34	**1.00**	0.67	0.00	0.00	0.36	0.24
#_trainings	0.67	0.00	0.67	**1.00**	0.00	0.25	0.41	0.21
#_meetings	0.00	0.00	0.00	0.00	**1.00**	0.38	0.25	0.12
#_calls	0.42	0.00	0.00	0.25	0.38	**1.00**	0.74	0.73
#_customers	0.34	0.12	0.36	0.41	0.25	0.74	**1.00**	0.58
#_bids	0.96	0.54	0.24	0.21	0.12	0.73	0.58	**1.00**

Table 3. Absolute correlation coefficient values of the preselected measures

value of A – which represents the sales managers' present-value monetary equivalent of complete information – we asked each sales manager how many daily rates he would pay for having complete information on cross-divisional cooperation during the planning horizon. We obtained an average of 10 daily rates, which we multiplied with the sales managers' average daily rate of 750 € and their overall number – there were 50. We finally obtained A = 375,000 €. As for informational disutility $D_{info}(x)$, the sales managers received sample reports. Each contained a different amount of measures, but had exactly the same layout as the reports that were planned to be finally used. The sales managers' task was to entirely understand the reports. For each amount of measures, we logged the time. In order to determine the value of B, we used the average value for 8 measures – which was 2.0 hours. We normalized it with respect to the sales managers' average daily working time – which was 9 hours. Then, we multiplied it with the sales managers' average daily rate and their overall number. As the report was planned to be presented monthly and the planning horizon was 3 years, we calculated the annual payments and the corresponding present value with an interest rate of 10 %. Assuming that the managers had to try to understand the report completely anew each time they received it, we obtained B = 273,554 €. We used the other time values for approximating β. The sales managers coped well with a low number of measures, but had problems with more than approx. 4–5 measures. So we obtained $\beta = 2.8$. The economic disutility $D_{econ}(x)$ was calculated based on Boehm's widespread cost estimation model CoCoMo (1981). Together with the company's DSS experts, we parameterized the estimation model as $PM=2.94 \cdot 0.20 \cdot LOC^{1.2}$ where PM and LOC denote person months and thousand lines of code respectively. The present-value effort for customizing a report with one measure and maintaining it during the planning horizon was estimated equivalent to 1.250 thousand lines of code. With the DSS experts' average daily rate of 400 € and 20 working days per month, we obtained C = 74,553 € and $\gamma = 1.2$.

On this basis, we determined the optimal fraction of measures $x^{opt}_{info+econ}$ by computing the joint informational and economic net utility $U_{info+econ,net}(x)$ (see Table 2 and *Figure 3*). If we had only considered the informational perspective, the highest informational net utility would have resulted from $x^{opt}_{info} = 0.5$. We would have selected the 4 measures with the highest informational utility, i. e. *%_leads, Ø_T_leads*, *#_meetings*, and *#_customers*. As we also took on an economic perspective, the highest joint net utility resulted from $x^{opt}_{info+econ} = 0.375$. We selected the 3 measures with the highest informational utility, i. e. *%_leads*, *#_customers*, and *#_bids*. Here, it becomes obvious that measures can be discarded if the number of selected measures changes.

We applied the same procedure to all other CSF. Advantageously, this caused considerably less effort because $D_{info}(x)$ and $D_{econ}(x)$ needed not to be determined anew. $U_{info}(x)$ could be calculated based on

No. of selected measures	0	1	2	3	4	5	6	7	8
Fraction x	0	0.125	0.25	0.375	0.5	0.625	0.75	0.875	1
$U_{info}(x)$ [€]	0	217,969	281,719	310,781	337,500	352,969	365,156	373,125	375,000
$D_{info}(x)$ [€]	0	810	5,640	17,552	39,279	73,368	122,240	188,220	273,554
$U_{info,net}(x)$ [€]	0	217,159	276,079	293,229	**298,221**	279,601	242,916	184,905	101,446
$D_{econ}(x)$ [€]	0	6,148	14,125	22,978	32,451	42,415	52,789	63,515	74,553
$U_{info+econ,net}(x)$ [€]	0	211,011	261,954	**270,251**	265,770	237,185	190,127	121,390	26,893

Table 2. (Dis-)Utility values for the CSF "cross-divisional cooperation"

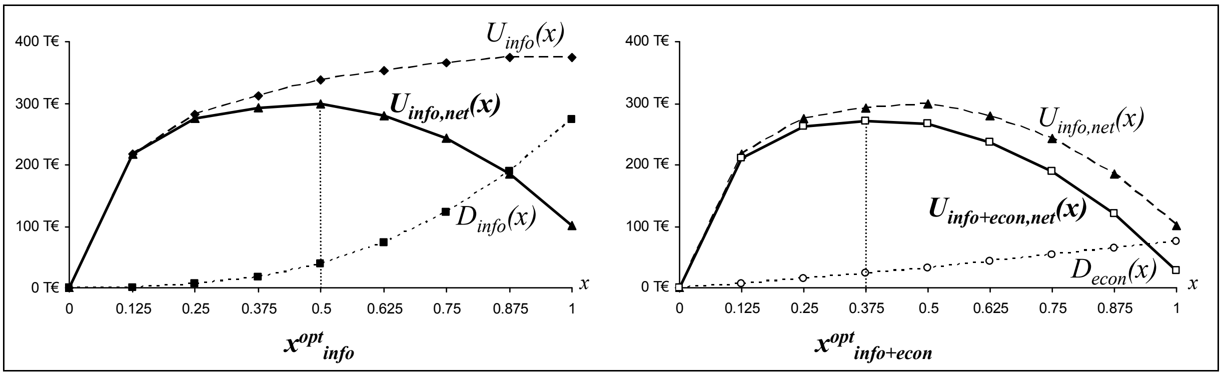

Figure 3. Visualization of the (dis-)utility values for the CSF "cross-divisional cooperation"

historical data so that the sales managers' expertise was only required for preselecting thematically appropriate measures and identifying meaningfully interpretable correlations. In sum, we modernized the company's sales reporting by identifying CSF, integrating non-monetary and leading measures, and significantly reducing the overall number of measures.

4.2 Checking the optimization model against the PMS requirements

The optimization model particularly aims at closing the primary research gap with respect to clarity (*R.2*) and interdependencies (*R.5*) (see section 2). Requiring to select a manageable amount of measures, clarity is addressed explicitly. Informational disutility expresses the decision makers' and reporting tools' ability of coping with information complexity and is contrasted to information utility. Thereby, we make sure that the whole amount of information does not become too complex and remains manageable for the decision makers. Interdependencies are also addressed explicitly. This is because informational utility uses interdependencies in order to express how "much" information measures provide about one another. The example showed – albeit in a simplifying manner – how an interdependency-based informational utility can be operationalized for non-monetary and leading measures with Pearson's correlation coefficient. Moreover, the model makes the process of measure selection more intersubjectively comprehensible (*R.4*). Although most parameters cannot be determined without subjective influences, just the fact that it is clear how they are formally linked increases intersubjectivity. Decision makers do not only participate by means of explorative interviews, but also in a structured manner by validating interdependencies and estimating the model's parameters (*R.6*). Concluding, the model does not only address the primary research gap, but also ameliorates the other requirements.

5 SUMMARY AND FUTURE RESEARCH

An optimization model has been proposed that helps to determine which and how many measures should be selected from a set of thematically appropriate measures in order to monitor specific fields of action (CSF). Both informational and economic objectives are considered. That is, (statistic) interdependencies among measures, decision makers' and reporting tools' ability of coping with information complexity as well as payments for customizing and maintaining reporting tools influence measure selection. The model's applicability was demonstrated with an illustrative example from a global electronics and electrical engineering company. The model will be subject to future research:

1. The optimization model is applied to one field of action a time. Several fields of action can only be addressed successively and isolated. The fact that measures may be thematically appropriate for more than one field of action is not considered. Hence, an integrated perspective is desirable and should be added.

2. So far, only measures from existing application systems are considered. On the one hand, this is reasonable as in many companies more measures exist than any decision maker can ever analyse. On the other hand, positive effects of innovative measures are neglected and need to be integrated.
3. (Data-driven) DSS and data warehouses do not only comprise measures, but also other master data for evaluation. In order to deal with their full scope, the model needs to be complemented e. g. by an approach which assesses relevant dimensions and dimension elements.
4. Although the model has been employed successfully in business practice, empirical evidence is missing with respect to whether its recommendations actually improve decision quality. It would be insightful and strengthen evaluation to conduct respective empirical studies.

References

Ackoff, R. (1967). Management Misinformation Systems. Management Science, 4 (14), B147-B156.

Alter, S. (1980). Decision Support Systems: Current Practice and Continuing Challenge. Addison-Wesley, Reading.

Artley, W. and Stroh S. (2001). The Performance-based Management Handbook (Volume 2).

Axson, D. (2007). Best Practices in Planning and Performance Management: From Data to Decisions. 2nd Edition. John Wiley & Sons, Hoboken.

Boehm, B.W. (1981). Software Engineering Economics. Prentice Hall PTR, Upper Saddle River.

Browne, G. and Ramesh, V. (2002). Improving information requirements determination: a cognitive perspective. Information & Management, 39 (8), 625-645.

Caplice, C. and Sheffi, Y. (1995). A Review and Evaluation of Logistics Performance Measurement Systems. The international Journal of Logistics Management, 6 (1), 61-74.

Davis, G. (1982). Strategies for information requirements determination. IBM Systems Journal, 21 (1), 4-30.

Duncan, J. (1980). The Demonstration of Capacity Limitation. Cognitive Psychology, 12 (1), 75-96.

Eccles, R. (1991). The Performance Measurement Manifesto. Harvard Business Review, 69 (1), 131-137.

Edwards, A. (1976). An Introduction to Linear Regression and Correlation. Freeman, San Francisco.

Evans, J. (2004). An exploratory study of performance measurement systems and relationships with performance results. Journal of Operations Management, 22 (3), 219-232.

Giorgini, P., Rizzi, S. and Garzetti, M. (2008). GRAnD: A goal-oriented approach to requirement analysis in data warehouses. Decision Support Systems, 45 (1), 4-21.

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

Inmon, W. (2005). Building the Data Warehouse. 4th Edition. Wiley & Sons, New York.

Küpper, H. (2005). Controlling: Konzeption, Aufgaben, Instrumente. 4th Edition. Schäffer-Poeschel, Stuttgart.

Leidecker, J. and Bruno, A. (1984). Identifying and Using Critical Success Factors. Long Rage Planning, 17 (1), 23-32.

Liebetruth, T. and Otto, A. (2006). Ein formales Modell zur Auswahl von Kennzahlen. Controlling, 18 (1), 13-23.

Miller, G. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. The Psychological Review, 63 (2), 81-97.

Neely, A., Mills, J., Platts, K., Richards, H., Gregory, M., Bourne, M. and Kennerley, M. (2000). Performance measurement system design: developing and testing a process-based approach. International Journal of Operations and Production Management, 20 (9/10), 1119-1145.

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

NIST (2008). Baldrige National Quality Program, Criteria for Performance Excellence.

Power, D. (2002). Decision Support Systems: Concepts and Resources for Managers. Quorum Books, Westport.

Rockart, J. (1979). Chief executives define their own data needs. Harvard Business Review, 57 (2), 81-93.

Sisfontes-Monge, M. (2007). CPM and Balanced Scorecard with SAP. Galileo, Bonn.

Watson, H. and Frolick, M. (1993). Determining Information Requirements for an EIS. MIS Quarterly, 17 (3), 255-269.