The development of a dynamic performance measurement system

Implemented in an India based IT outsourcing company

Patrick Alexander Bonouvrie

11/11/2011

Studentnr.: s0166928

Commissioned by: '01 Synergy', Ludhiana, Punjab, India.

Contact Persons: Preet Chandhokee, Davinder Singh

Examiners: A.B.J.M Wijnhoven, C. Amrit

NEDERLANDSE SAMENVATTING

De huidige methodes voor het ontwikkelen van performance measurement systems voldoen niet aan de eisen van een performance measurement systeem van vandaag. De huidige methodes verliezen kort na implementatie haar relevantie omdat ze niet effectief kunnen omgaan met een snel veranderende bedrijfsomgeving. Bovendien bieden deze systemen vooral een nieuwe bron van data voor de managers, maar bieden zij niet voldoende de mogelijkheid om ook als hulpmiddel in het beslissingsproces te dienen.

Door verschillende beslissings- en simulatie modellen toe te passen op de huidige statische methodes, kan een dynamisch performance measurement systeem ontworpen worden die wel voldoet aan de eisen van vandaag en niet haar relevantie verliest bij wijzigingen in de bedrijfsomgeving. Aan de hand van een case studie bij een IT outsourcing bedrijf in India, wordt een model van zes stappen uitgelegd waarmee zo'n dynamisch performance measurement systeem ontwikkeld kan worden. Dit systeem zal snel aangepast kunnen worden als er om aanpassingen gevraagd wordt. Tevens zal het systeem niet alleen een nieuwe bron van data zijn voor de managers, maar tegelijk ook een mogelijkheid bieden om beslissingen te maken aan de hand van deze data door simulaties uit te voeren waarmee mogelijke effecten van gekozen policy's gesimuleerd kunnen worden.

De zes stappen van het model zien er als volgt uit;

1. Stap 1: Ontwikkel key business objectives vanuit de strategie aan de hand van de Balanced Scorecard methodiek.
2. Stap 2: Genereer attributen waarmee de in stap 1 gevonden key business objectives meetbaar gemaakt kunnen worden.
3. Stap 3: Test de gevonden attributen op betrouwbaarheid en validiteit, en documenteer deze attributen.
4. Stap 4: Interpretatie van de attributen;
    a. Pas een beslissingsmethodiek toe op de attributen om haar gewichten te bepalen.
5. Stap 5: Ondersteuning van het systeem bij het maken van beslissingen;
    a. Bepaal causale relaties tussen de attributen en tussen de key business objectives.
    b. Genereer de simulatie formule;
        i. Deel 1) Bepaal de coëfficiënten van de invloed van verschillende attributen op de key business objectives aan de hand de bij Stap 4 bepaalde gewichten.
        ii. Deel 2) Bepaal de onderlinge groei coëfficiënten van de verschillende attributen.
6. Stap 6: Evaluatie van het systeem.
    a. Test het systeem in de praktijk, breng wijzigen aan waar nodig.

Table of Contents

ABSTRACT

The current set of performance measurement frameworks does not satisfy today's requirements of a performance measurement system. Such systems loose relevance and effectiveness in today's quickly changing environments due to the lack of modifiability and the lack of showing the impact of decision making. By applying decision and simulation models such as MCDA and System Dynamics to static performance measurement frameworks, a dynamic performance measurement system can be created which is able to maintain its effectiveness in today's organizations. A six step approach based on experiences at an Indian based IT Outsourcing company is suggested in order to create such a system, resulting in a performance measurement framework that can quickly be modified and not only generated historic data but is also able to simulate the impact of possible decisions based on the data, allowing the executives to choose the best policy.

1. INTRODUCTION & BACKGROUND

The purpose of this report is to explain a method for the design of a dynamic performance measurement system. This study is based on experiences at the headquarters - which is also the main development lab - of an India based IT outsourcing company called '01 Synergy' located in the city Ludhiana, state Punjab, country India.

'01 Synergy' offers a wide variety of services to a client base mainly located in Western Europe and North America including Global 500 companies. '01 Synergy' has development IT labs in Pune, Delhi and Ludhiana (India), and offices in Canada and the USA. '01 Synergy' deploys around 180 IT engineers.

The company operates in a rapidly changing environment and suffers from problems how to control her productivity. The high variety of clients leads to a high variety of projects and this results in roles and responsibilities changing quickly. '01 Synergy' actively recruits from technical universities in the regions of their development labs and aims to offer advanced training programs to undergraduates. Employee's task statements often depend on the nature of the current projects, leading to problems with performance evaluation. The high variability of the tasks of some of the teams requires a performance measurement system that is able to quickly adapt to environmental changes. Part of the salary of the employees is based on incentives. Currently, these incentives are calculated with help of out-dated performance evaluation sheets. These performance sheets do not represent the current work accomplishments well, nor for the teams with a high variability of tasks, nor for teams with a low variability of tasks.

Several authors including Kaplan & Norton (2001) and Bourne, Neely, Mills & Platts (2003) state that in order to be successful in quickly changing environments the company should align the organization to a clear strategy. [1, 2] A Performance Measurement System (PMS) offers a method to translate the company's strategy into daily operations. Most of the research on PMS focuses on manufacturing industries, and just a few on service industries. '01 Synergy' offers consulting services (mainly business process outsourcing) and manufactures products (mobile-/software-/web-developments). This means that the organization requires a performance measurement system which is able to address both the consulting and the manufacturing teams. Finally, in order to be able to adapt to the quickly changing environment, it is required that changes can be made to the PMS continuously.

After identifying the key problem by showing causes and relations, a literature review on PMS will be conducted. The review will elaborate on the PMS in global, and will present methods on how to add the dynamic character to the PMS. When the review has been done, a methodology will be conducted and the development of the PMS will be presented. Finally an example of the output and suggestions for further research will be given.

2. PROBLEM IDENTIFICATION

After observing the daily routines and gathering company data through interviews, the main problems and its causes were identified. Overall the key problems seemed to be the lack of available resources and the inefficient personal development programs.

Currently the work pressure is high due to a lack of available resources. The pressure is particularly high for the single technical executive, who carries all responsibilities that require technical knowledge. This results in him being assigned too much work and too many responsibilities. In order to compensate the lack of resources, trainees are asked to interrupt their training programs in order to assist with current projects. Since they lack both professional experience and required knowledge, they mainly assist with basic tasks such as data entry. On this short term, this reduces the need of resources, but on the long term, however, this greatly hurts the productivity of the company. This leads to the second problem, inefficient personal development programs.

Whenever trainees interrupt their training programs future capital will be wasted. The trainees will not be able to develop their technical skills at the same rate as when they would spend their full time on their personal development. The actual gains for assisting at the projects are small for the trainees, the amount of knowledge they gather from the data entries is much smaller compared to executing their training plans. This leads to a vicious circle, the lack of resources remains due to new projects coming in, and the trainees cannot spend their full time on training themselves and thus their technical skills remain at a low level. This hurts the long term quality of the company recourses, the productivity will barely increase, the need of resources remains and thus the lack of training remains as well. It should also be noted that a lack of personal development may lead to a decrease of the satisfaction of the trainees, which may result in them leaving the company. Failing to train the trainees will hurt the long term quality of resources, but losing the trainees to another company will hurt even more; also wasting the time, money, and effort put into the training program.

Since the recruitment of experienced developers in the new development technologies appears to be a problem, personal development has a huge potential value to the company, both for trainees and for current employees. When personal growth is guaranteed, experience within the company will increase; trainees will be able to develop into full time programmers and thus decrease the need of resources. The work pressure of the technical executive will decrease and new trainees can continue their training programs without an interruption. These expectations result in the first strategy that should address the problems, ‘S1 – Personal Development’. Finally, with a more effective project management, projects should not be accepted (or at least outsourced) if the company cannot meet the required amount of resources, leading to a second strategy; ‘S2 – Project Management’.

The above obervations are mapped in the Causal Loop Diagram (CLD) in Figure 1. The diagram explains how the company currently behaves and shows the structure of the problems by displaying the causes and relations between several aspects. The CLD can be used to understand the consequences of certain actions and the interaction between several problems. This CLD is focused around Productivity, which is defined as the total power of the company to complete the current projects according to the client’s wishes. Also, the problems will be approached from the viewpoint of human capital.

The CLD consists of arrows marked with a '+' or a '-', which shows a positive or a negative relation. A positive relation means an *increase* of X leads to an *increase* of Y, or a *decrease* of X leads to a *decrease* of Y. A negative relation means an *increase* of X leads to a *decrease* of Y, or a *decrease* of X leads to an *increase* of Y. There are two types of causal loops, the 'N' type and the 'P' type. The 'N' type contains the 'negative feedback' loops and is a 'balanced' loop, meaning an increase of X will lead to a certain set of actions which will lead to a decrease of X again and visa versa. A balanced loop has an uneven number of negative links. This type of loop thus balances out the effect of a change. The second type, 'P', contains the 'positive feedback' loops, also called 'reinforcing' loops, meaning that an increase of X will lead to a certain set of action which will lead to an even higher increase of X, or visa versa. A reinforcing loop has an even number of negative links and is associated with exponential increases or decreases. By changing a variable within the loop, the loop can be changed from a reinforcing loop to a balanced loop. [3]

The nodes in the figure behave as connecting points, this means that when node 'A' is reached, the path will continue at the other node 'A' in the model.

N1 and P1 correspond with the earlier described key problems, whereas the other loops are smaller problems explained in Annex 4.

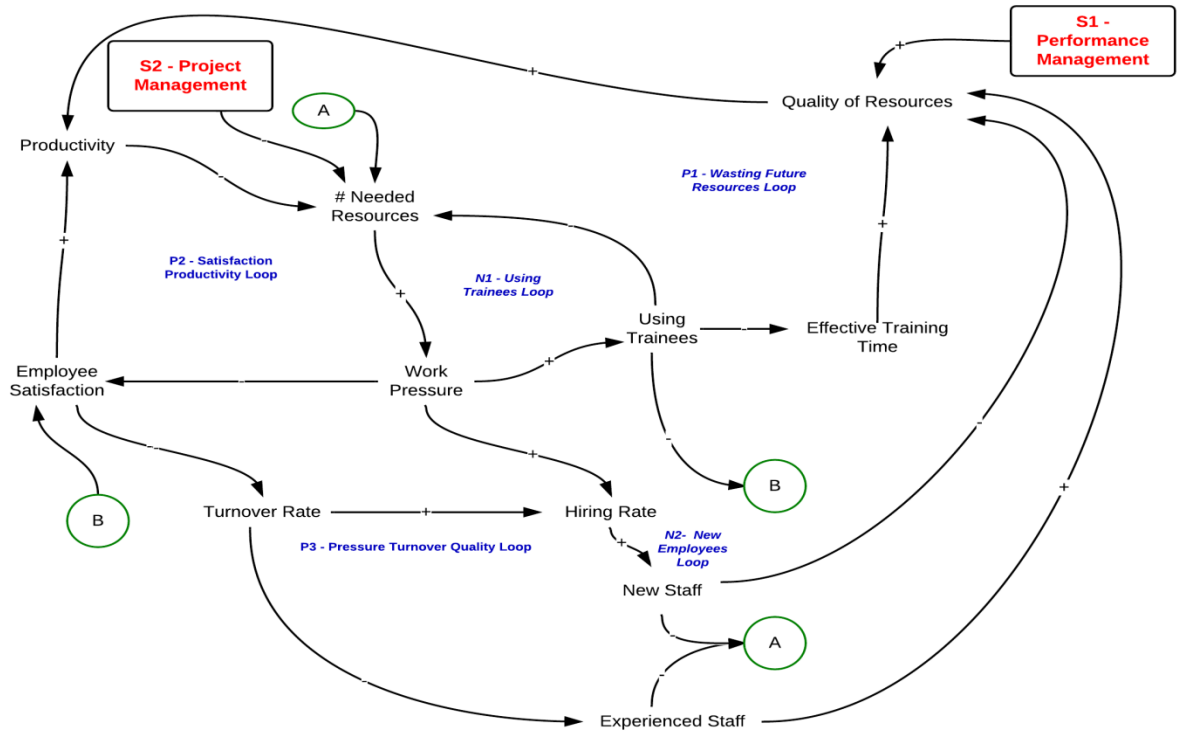

FIGURE 1:CLD PROBLEM IDENTIFICATION

By having an effective performance measurement system in place, the efficiency of many aspects of the company will increase. The employee's responsibilities will be clearer, benefiting both the on-floor efficiency and the recruitment process. Trainees will be offered training plans, which allows them to develop themselves into full-time programmers. The company's expectations of the employees will be more explicit, decreasing the uncertainty amongst employees, increasing the trustworthiness of the management and being the foundation of future personal development plans. [4] Due to this, higher gains were expected by executing the S1 – Personal Development compared to S2 – Project Management. Since the company

operates in a quickly changing IT outsourcing environment, the system should have a dynamic character in order to be effective. Compared to a regular PMS, a dynamic PMS should be able to be modified and the system should be able to simulate the impact of decision making based on the data generated by the system. This leads to the following leading question;

WHAT ARE THE REQUIRED PROCEDURES FOR INTRODUCING A PERFORMANCE MEASUREMENT SYSTEM AT '01 SYNERGY', THAT IS ABLE TO EVALUATE BOTH HIGH VARIETY AND LOW VARIETY FUNCTIONS AND HOW CAN THIS SYSTEM COMPLY WITH BOTH MODIFIABILITY AND SIMULATIONS OF DECISION MAKING?

The main question can be answered by the following sub questions;

What kind of performance measurement system applies best to '01 Synergy'?

What are the requirements in order to create an effective performance measurement system?

What are the advantages of a dynamic performance measurement system compared to a regular performance measurement system?

How to guarantee the dynamic character of the performance measurement system?

How to implement and maintain a dynamic performance measurement system?

3. LITERATURE REVIEW

The need for and design of performance measurement systems will be explained first, followed by ideas how to change from static performance measurement systems to dynamic measurement systems. We will also discuss how to calculate effective measures, how to test the reliability and validity of the measures and how to increase the chance of a successful implementation of the performance measurement system.

3.1 THE NEED FOR DYNAMIC PERFORMANCE MANAGEMENT SYSTEMS

The business environment of today is changing rapidly and the complexity of the environment of the organizations increases. In order to survive, organizations need to adapt accordingly. In this rapidly changing climate, it has never been more important to implement solid strategies. Research, however, shows that companies struggle in executing the strategies needed to stay competitive. [5]

According to Kaplan & Norton (1996), one of the reasons for this "is clearly that while these strategies, and the business issues behind them, are changing constantly, the tools for measuring the effectiveness of these strategies have not kept pace" (p. 2). The traditional performance management tools originated from financial accounting measures that were introduced within companies after the First World War, such as variance analysis, standard costing and return on investment. In the following fifty years until around 1980, however, there were no significant developments of these accounting measures. Several authors started to criticize these traditional measures, which are still being used in businesses today, for having a narrow, one-dimensional focus. Other critics include;

- The encouragement of short-termism, for example the delay of capital investment.
- The lack of strategic focus.
- The fail to provide data on quality, responsiveness and flexibility
- The encouragement of local optimization – for example "manufacturing" inventory to keep people and machines busy.
- The encouragement of managers to minimize the variances from standard rather than seek to improve continually.
- The fail to provide information on what customers want and how competitors are performing. [2]

As a consequence of these critics, the interest in developing a balanced performance measurement system increased during the early 1990s, resulting in the creation of frameworks such as 'The Balanced Scorecard (BSC)', the 'Performance Pyramid' and the 'Results and Determinants Matrix'. [2] Compared to other performance frameworks, the balanced scorecard provides an excellent balanced structured framework for aligning the performance management system to the organization's strategy. According to Hudson, Smart, & Bourne (2001) the main problem of the performance pyramid is that this framework fails to specify "either the form of the measures or the process for developing them." (p. 1103), whereas the determinants matrix "does not include customers or human resources as dimensions of performance and cannot, therefore, give a truly balanced view of performance." (p. 1104). Other newer models, such as the 'Integrated PM system methodology' and the 'Cambridge PM process' also offer a framework that covers most of the performance measurement criteria found by

Hudson *et al.* (2001), however they lack the structure for designing the process.[6]. Due to above stated reasons, the balance scorecard will be used to translate the companies' strategy into measurable attributes.

After the performance framework has been created, the system should be used to manage the performance of the organization. The most heard of disadvantages of the balanced scorecard are that it fails to maintain the relevance of the measures, and that it fails to specify a user-centered development process. [7] The user-centered process can be achieved by designing the found measures specifically and defining clear objectives and targets. After all, if the implementation of the Balanced Scorecard is successful, the organization will move in the direction of a learning organization, after which the new culture will create an internal environment of continuous improvement and personal development.[8] However, Santos, Belton, & Howick (2001) state;

> What seems to happen with the existing PMS is that they tend to provide a large and complex amount of information about the performance of the organi[z]ation and whether corrective actions are required or not. However, these systems neither provide participants with tools to assist decision makers understand, organi[z]e and use such information, in order to identify for example the causes of poor performance, nor provide participants with tools to help them in evaluating and eventually selecting appropriate corrective actions. One of the most common complaints made by practitioners is that PMS provide too much data and too little analysis. [8]

Furthermore, as Akkermans & Oorschot (2005) experienced in their case study, executers of the performance measurement system often doubt the quality of the found performance indicators. By applying decision and feedback models, the issues raised by Akkermans *et al.* and Santos *et al,* can mostly be addressed. In order to analyze and continuously improve the measures by creating a feedback system, System Dynamics (SD) can be used. System dynamics models simulate how different aspects of a system interact with each other in order to map the behavior of the system over time. These simulations are often used as policy analysis tools to show consequences and connections. In order to evaluate the outcome and in order to support decision makers using these models, multi-criteria decision analysis (MCDA) can be used. [9] MCDA will be applied to assist the decision makers in their interpretation of the data generated by the PMS, whereas SD will help the decision makers to choose the best policy. [3, 8, 9]

Both Santos *et al.* (2001) and Akkermans & Oorschot (2005) suggest a two-step approach in applying system dynamics to performance measurement systems improvement. During the first step, the qualitative one, Causal Loop Diagramming (CLD) will be applied to create a strategy map by showing the relations between several measures. A CLD "gives a clear picture of the different elements of the problem and the interconnectedness between them (cause and effect, feedback loops, delays and so on). […] Notice that the use of CLDs allows to identify feedback loops, and it is the interaction between these loops that determines the dynamics of the system." [8] Using CLD to identify and structure performance measures offers various advantages. It ensures the measures were designed in line with the strategy of the organization. Furthermore the strategy map shows if the found measures encourage correct strategic behavior. Also, if objectives change, it directly shows which measures are connected with the objective and thus should be adjusted as well. As an extra benefit, this also leads to people reviewing and clarifying their objectives leading to an increased insight in the situation. Finally the model provides the basis for future analysis. [8] In the second step a quantitative SD simulation model is designed.

This model is based on the CLD that has been created in step one. The model is essential when testing and comparing different courses of action to increase organizational performance. By applying the company data, a graph that partly replicates history and partly predicts future behavior can be generated. [3, 10, 11]

Multi-Criteria Decision Analysis (MCDA) is a technique that assist decision makers in their decision making process. MCDA allows making decisions based on multiple criteria. Even though the technique allows multiple criteria to be weighted into the final verdict, the preferences of the decision maker are still clearly reflected into the results of each MCDA technique. Many different methods of MCDA can be found in the literature, such as AHP, MAUT and SMART.[12] All methods have different grades in complexity and accessibility, but in all methods the decision makers' preferences are reflected. SMART is a simple multi attribute weighting method based on ratio estimation. As Mustajoki, Hamalainen, & Salo, (2005) state "…the true usefulness of the methods is determined by procedural aspects." [12] SMART is an easy-to-use approach compared to AHP and MAUT. Since the decision makers who will use the PMS will have to be able to easily and quickly change parameters within the PMS in order to adapt to the ever changing environment, SMART is preferred above the more complex AHP and MAUT. In SMART the decision maker is asked to identify the most important attribute, and assign this attribute a value of hundred. After identifying the particular attributes, the decision maker is asked to assign a value to each other attribute to denote the relative importance compared to the most important attribute. The actual weights are then being determined by normalizing the sum of the given values. [12, 13]

3.2 THE BALANCED SCORECARD

Initially, Kaplan & Norton (1992) defined the balanced scorecard as a framework that "provides a medium to translate the [company's] vision into a clear set of objectives. These objectives are then further translated into a system of performance measurement that effectively communicates a powerful, forward-looking, strategic focus to the entire organization." [14] Its aim was to design the key success factors of an organization and to align the daily routines to the strategy of the company. Kaplan & Norton believe that financial results are achieved by the alignment and implementation of strategy, instead of being their driving force as traditional measures suggest. [14]

The original balanced scorecard, as shown in Figure 2 features four perspectives; the customer perspective, the financial perspective, the internal-business-process perspective, and the learning-and-growth perspective. The scorecard ensures an overall view of the organization by covering three of the major stakeholders (customers, employees, shareholders) within these perspectives. The measures chosen within the perspectives should be derived from the strategy of the company. The financial perspective includes strategies for growth, profitability, and risk, viewed from the perspective of the shareholder. The customer perspective includes strategies for creating value and differentiation from the perspective of the customer. The internal-business-process perspective includes strategic priorities for the critical internal processes in which the organization should excel, creating customer and shareholder satisfaction. Finally, the learning-and-growth perspective includes the strategic priorities to create a climate that supports organizational change, innovation and growth. [1]

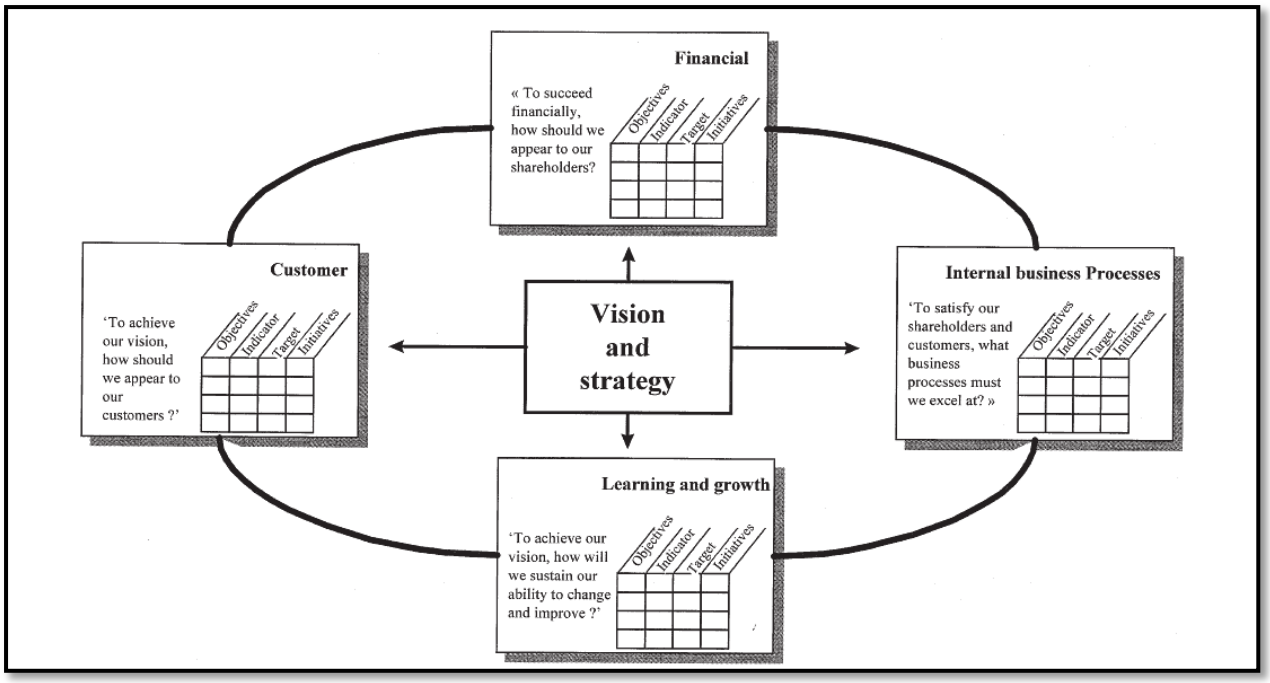

FIGURE 2: THE BALANCED SCORECARD ((**ERROR! REFERENCE SOURCE NOT FOUND.**)KAPLAN AND NORTON (1996A, P.197))

The actual function of the balanced scorecard differs per organization, depending on the goals the organization aims to achieve. These goals can range from gathering data to question the current strategy, creating the environment for 360^0 feedback process, to being the key part of the whole management system. An overview of the functions is shown in Figure 3. The balanced scorecard is able to cover all four main areas of the management system, however mostly one or two sections will dominate when implementing the scorecard, depending on the aims of the organization. [7]

Organizations often have different management systems in place. These systems all initiate a particular behavior of the employees. However, most of the systems are standalone system, all with their own purpose. They lack the integration with the other systems, leading to a lack of overview of the whole situation. By substituting the separate systems by the balanced scorecard, the different systems can be integrated and aligned to the company's strategy. [7]

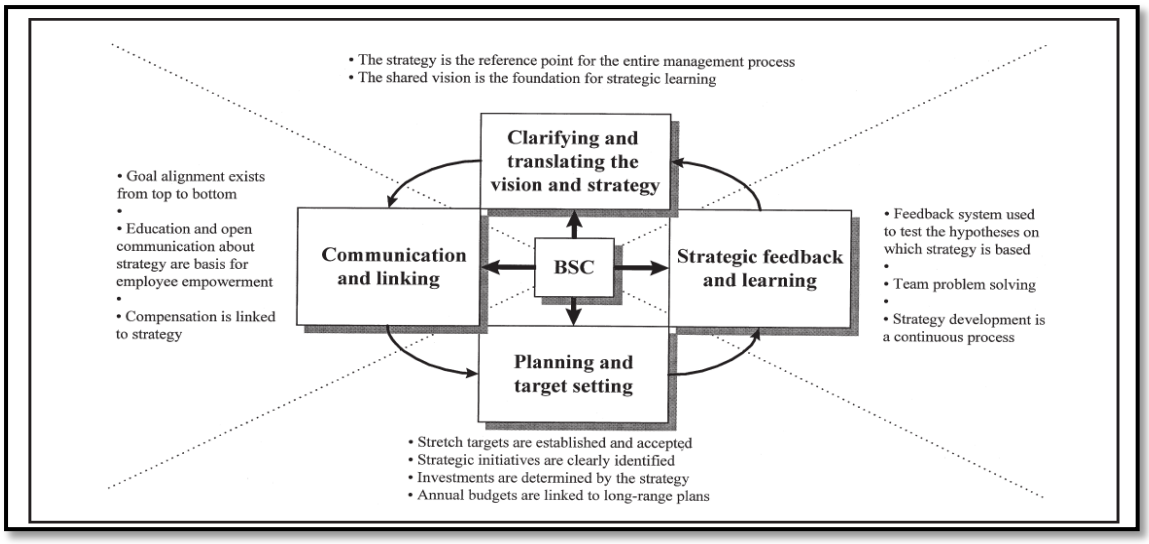

FIGURE 3: A MANAGEMENT SYSTEM FOR STRATEGIC IMPLEMENTATION [KAPLAN AND NORTON (1996A, P. 197))

3.3 PERFORMANCE MEASUREMENT IMPLEMENTATION

Bourne, Neely, Platts, & Mills (2002) identify some of the issues managers experience by designing and implementing performance measurement systems. First, four critical factors in the process of development of the performance system were found; point of entry (how the introduction and launch was handled), participation (who was involved), project management and procedure (the tools used in the process itself). However, successfully handling these critical factors may not be sufficient for the successful implementation of the system. Other non-process factors should also be valued. Secondly, Bourne *et al.* (2002) state that the majority of the implementation problems named in the literature are caused by bad design. Thirdly, they identify four mayor project specific implementation blockers; the required effort for implementation, the consequences of performance measurement, priority shifts to other initiatives and the easy of data accessibility. Finally, top management commitment (and their perceived benefits) is of crucial importance to the success of implementation of the performance measurement system. It is said that management commitment will mostly not be static, but fluctuates during the process. [15]

Measures should be tested for reliability and validity. Validity is defined as "..the extent to which any measure measures what it is intended to measure"[16], whereas "..consistency found in repeated measurements of the same phenomenon is referred to as reliability." [16] Reliability of the found attributes will be determined by using test-retest. For determining the coefficient for this kind of reliability the standard error of measurement as a coefficient of variation (CV) will be used. CV is the standard deviation expressed as a percentage of the mean. [16] Multiple categories of validity are known, of which criterion-related validity is the most relevant in our situation. As Carmines and Zelle (1979) state, criterion related validity ".. is at issue when the purpose is to use an instrument to estimate some important form of behavior that is external to the measuring instrument itself, the latter being referred to as criterion."[16] The degree of the validity depends on the validity coefficient, which is the correlation between the criterion and its test.

According to Neely *et al.* (1997) the development of performance measures is a complex process. They argue that performance measures should include not just the formula, but also "the purpose of the measure, the frequency of measurement and the source of data all have to be considered" [17]. Furthermore, they state that "… inadequately designed performance measures can result in dysfunctional behavior. Often because the method of calculating performance – the formula – encourages individuals to pursue inappropriate courses of action." [17] People adapt to performance measures in order to ensure a positive outcome, even if this results in taking a course of action that hinders the positive results of the process. Thus, the designers of performance measures should mind the possible behavioral outcomes of each measure before implementing them. The measures should encourage the desired behavior. [17]

Neely *et al.* (1997) present a performance measure record sheet (Annex 2) that should lead to the design of good measures. The sheet includes all recommendations as shown in Figure 4: Performance Recommendations[17], except for two groups of measures. The first group includes measure 5; i.e. both the supplier and customer should be involved in the definition of the measure, measure 12; performance measures should be consistent, measure 17; performance measures should use data which are automatically collected as part of a process whenever possible, and measure 18; performance measures should be reported in a simple consistent format. The measures in this group are important process guidelines instead of actual measure design guideline, and complement the total framework next to the performance measure record sheet. The second group includes measure 10; performance measures should have visual impact, measure 11; performance measures should focus on improvement not variance, measure 16; performance measures should employ ratios rather than absolute numbers, measure 19; performance measures should be based on trends rather than snapshots, and measure 22; and performance measures should be objective not based on opinion. This group requires further research since "only anecdotal evidence exists to support [these] assertions" (p.49). [17]

	Recommendation
1	Performance measures should be derived from strategy
2	Performance measures should be simple to understand
3	Performance measures should provide timely and accurate feedback
4	Performance measures should be based on quantities that can be influenced, or controlled, by the user alone or in co-operation with others
5	Performance measures should reflect the "business process" – i.e. both the supplier and customer should be involved in the definition of the measure
6	Performance measures should relate to specific goals (targets)
7	Performance measures should be relevant
8	Performance measures should be part of a closed management loop
9	Performance measures should be clearly defined
10	Performance measures should have visual impact
11	Performance measures should focus on improvement
12	Performance measures should be consistent (in that they maintain their significance as time goes by)
13	Performance measures should provide fast feedback
14	Performance measures should have an explicit purpose
15	Performance measures should be based on an explicitly defined formula and source of data
16	Performance measures should employ ratios rather than absolute numbers
17	Performance measures should use data which are automatically collected as part of a process whenever possible
18	Performance measures should be reported in a simple, consistent format
19	Performance measures should be based on trends rather than snapshots
20	Performance measures should provide information
21	Performance measures should be precise – be exact about what is being measured
22	Performance measures should be objective – not based on opinion

FIGURE 4: PERFORMANCE RECOMMENDATIONS (NEELY ET AL. (1997))

According to Neely *et al.* (1997) the PMS should encourage an ideal behavior amongst employees, a behavior that will contribute to the overall business strategy. During the design of every single measure, the expected behavior for implementing such a measure should be kept in mind. [17] Furthermore, in order to ensure a high probability of successful implementation of a performance management system within an organization, Bourne *et al.* (2002) acknowledge the importance of a well-designed approach. [15]

The found constraints should be the backbone during every single step of the design and implementation process of the PMS. One of the preconditions of the PMS is that the system should be able to adapt to expected changes in the environment. This also means, indirectly, that the system should be easily understood and updated by the HR executives. Thus during the creation of the system it should be kept in mind that the PMS should be mostly automated and accessible, otherwise the PMS will most probably not be accepted within the organization. Finally, it should not be forget that the creation of performance measures is an iterative process. This process does not end after the design of the first set of measures, but demands continuous reviewing and improvement, as shown in Figure 5. [8] During the design phase, SD and BSC will be applied, the interpretation of the data when using MCDA will happen during the measure phase, afterwards the results will be analyzed and planned with SD.

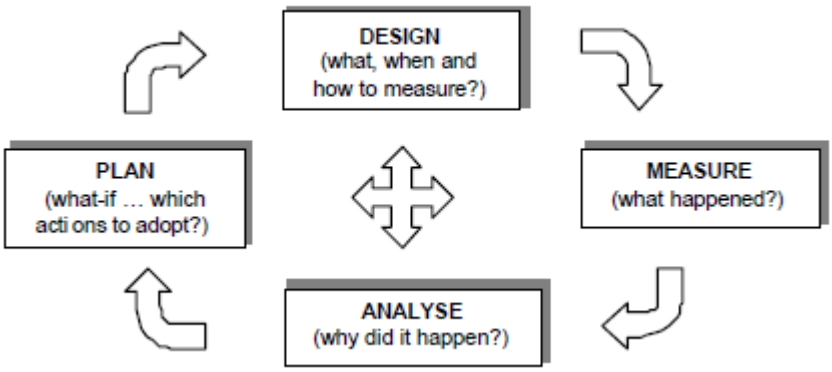

FIGURE 5: CONTINUES CIRCLE (SANTOS ET AL. (2001))

3.4 THE ADVANTAGE OF A DYNAMIC SYSTEM

Much has been written about the need and design of performance measurement systems, however, even though several authors such as Kaplan & Norton mention that performance measurement systems help to move the organization into the direction of a learning organization[1], little attention is given to how to modify the system in order to ensure its effectiveness when the environment changes. Furthermore, most suggestions of the authors are limited to static measures or systems based on historical data. Finally, little was written about the application of performance measurement systems to decision making. Performance measurement system should be dynamic; i.e. it should be modifiable whenever the environment changes and the impact of decision making based on the generated system-data should be mapped. As Kennerley, M. and A. Neely (2002) state ".. Consideration is being given to what should be measured today, but little attention is being paid to the question of what should be measured tomorrow. .." [18].

A static performance measurement system is a valuable tool at the exact moment the system is released, but by using a static system the relevance and thus the effectiveness of the system will quickly decrease in most of today's environments. A static performance measurement system may still be applicable in extremely stable industries, however, a dynamic system should be chosen in most cases.

4. PERFORMANCE MEASUREMENT SYSTEM DESIGN

The aim of the performance measurement system is to measure the performance of the employees, rather than the actual manufacturing. In a quickly changing environment as observed at '01 Synergy', it is of high importance to align the daily operations to the organization's strategy.[14]. This chapter will discuss the development phases needed to create a dynamic performance measurement system.

Peffers, Tuunanen, Rothenberger, & Chatterjee, (2008) suggest a method where designers can approach design problems systematically. They state a design methodology would include three elements, 1) conceptual principles to define what is mean by design science research, 2) practice rules, and 3) a process for carrying out and presenting the results.[19]. Peffers *et al.* (2008) developed a design process model as shown in Figure 6. In our case we enter the process model at the 'Problem Centered Initiation' stage. The first two activites of the model have been described in the first chapter.

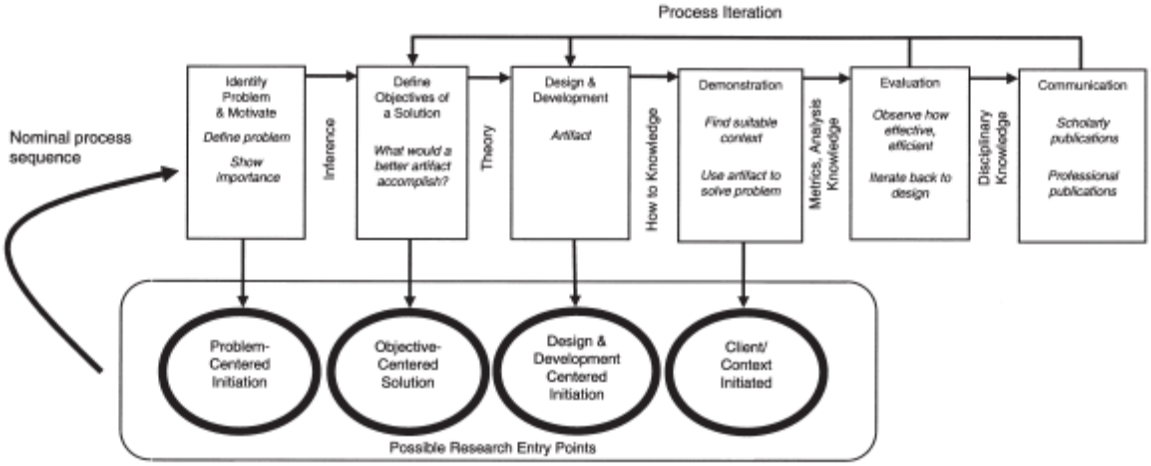

FIGURE 6: DSRM PROCESS MODEL

The design & development of the actual performance measurement system will be split into six phases. The chosen approach is based on the balanced scorecard principle of Kaplan & Norton (1992). While covering four major fields of the organization, i.e. financial, customer, learning & growth and internal business process the daily operations should be aligned to the company's vision and strategy. During the first phase key business objectives should be derived out of the strategy of the company, while at the second phase these business objectives should be translated into measurable attributes. Then in the third phase the attributes will be tested for reliability and validity, and documented in excel sheets. Phase four describes the interpretation of the attributes by applied swing weights to the attributes. These four steps combined result in the first draft of the PMS. System dynamics will be applied in Phase five to assist the executers of the PMS with the decision making; the relationship between the found attributes will be mapped into a causal loop diagram, following the process suggested by Akkermans & Oorschot (2005) in order to increase insight in the dynamics of the model. The goal of mapping the relations between the attributes is to give the executives an overview of the causes and expected results of certain policies. In order to understand the behavior of the organization, the management should be able to understand the consequences of the executed policies and the

overall relations between the attributes. Finally, in step 6, the system will be evaluated. An overview of the above process is shown in Figure 7: Basic Overview Methodology.

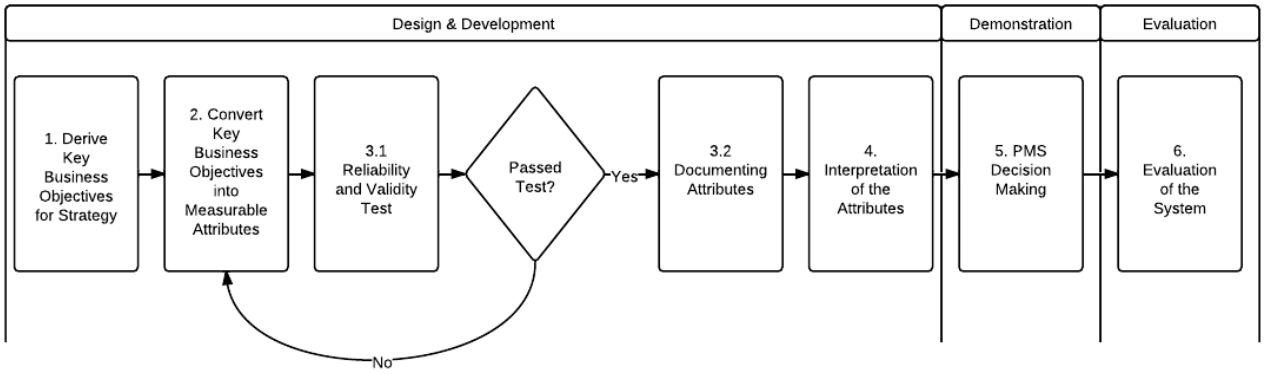

FIGURE 7: BASIC OVERVIEW METHODOLOGY

The separate phases of the design of the performance measurement system are explained in the tables below. The main theories used in every phase are explained in the left column, the key parts of the process are explained in the right column.

PHASE 1: DERIVE KEY BUSINESS OBJECTIVES FROM STRATEGY

TABLE 2: PHASE 1: DERIVE KEY BUSINESS OBJECTIVES FROM STRATEGY

Theories Used	Data Collection
'Balanced Scorecard' [1, 14, 20, 21] **'Requirements for successful a PMS' [2, 6, 15, 17, 22]**	Data will be collected through various meetings with the CEO and CTO. At first it should be made clear that both the CTO and CEO have the same view on the company's vision/strategy. Then my understanding with the company's strategy should be matched with theirs. Finally a start can be made with deriving key business objectives from the strategy. This will be done with keeping the focus on the four perspectives as named by the BSC. *Note: By analyzing the strategy during the meetings, it may very well be that it appears strategy should be changed on some aspect. Be sure to only start deriving the objectives after there is a general consensus.*

PHASE 2: CONVERT KEY BUSINESS OBJECTIVES INTO MEASURABLE ATTRIBUTES

TABLE 3: PHASE 2: CONVERT KEY BUSINESS OBJECTIVES INTO MEASURABLE ATTRIBUTES

Theories Used	Data Collection
'Balanced Scorecard' [1, 14, 20, 21] **'Requirements**	The business objectives should be translated into measurable attributes with help of the GM and the HR Executive. Their help is of high importance due to their high 'on-floor' knowledge. Furthermore their involvement will greatly increase the chance of acceptance and thus successful implementation, as the GM and HR executive should execute and maintain

for successful a PMS' [2, 6, 15, 17, 22, 23] **'Designing Performance Measures' [2, 17]**	the PMS. The translation of the objectives into attributes is done through brainstorm sessions while focusing on the four perspectives of the BSC and keeping Neely *et al*. 's requirements for effective performance measures in mind.

PHASE 3: TESTING THE RELIABILITY AND VALIDITY OF THE ATTRIBUTES AND DOCUMENTATION OF THE ATTRIBUTES

TABLE 4: PHASE 3: TESTING THE RELIABILITY AND VALIDITY OF THE ATTRIBUTES AND DOCUMENTATION OF THE ATTRIBUTES

Theories Used	Testing & Documentation
'Requirements for successful a PMS' [2, 6, 15, 17, 22]	Reliability of the scorer will be determined by comparing the monthly means of the total scores given at each key objective. A scorer is reliable if the average change in mean will be below 10%. An attribute is considered reliable whenever the CV percentage is below 15%.
'Designing Performance Measures' [2, 17] **'Reliability and validity assessment' [16]**	Documentation of the found attributes using Neely *et al.* (1997) performance measure framework. Clearly define the attributes, find targets and expected behavior. Create overview excel sheets per function. *Note: At the used framework the 'who owns the measure?' and 'what do they do?' questions will be left out, since they are not relevant due to the small size of the office and because we are measuring the performance of people rather than production units. May the company want to expand the system to the other offices, these questions could become relevant.*

PHASE 4: INTERPRETATION OF THE ATTRIBUTES

TABLE 5: PHASE 4: INTERPRETATION OF THE ATTRIBUTES

Theories Used	Determining the weights.
Applying MCDA to PMS [8] **MCDA [12]**	Swing weights will be applied to add weights to the found attributes. The actual determination of the weights will be done in during a meeting with the GM and the HR, and will be send to the CEO and CTO for review.

PHASE 5: PERFORMANCE MEASUREMENT SYSTEM DECISION MAKING

TABLE 6: PHASE 5: PERFORMANCE MEASUREMENT SYSTEM DECISION MAKING

Theories Used	Testing Procedure
Applying SD to PMS [3, 8]	The cause and effects of the total set of attributes will be graphed in a causal loop diagram according to the insights of Santos *et al.* and Akkermans *et a*l. This CLD should support the management in their understanding of the system and in deciding which policies to execute and in understanding the expected effects of each policy.

PHASE 6: EVALUATION OF THE SYSTEM

TABLE 7: PHASE 6: EVALUATION OF THE SYSTEM

Theories Used	Testing Procedure
'Requirements for successful a PMS' [2, 6, 15, 17, 22] **'Designing Performance Measures' [2, 17]** **Evaluation [19, 24]**	After the first set of performance sheets are finished, they will be tested in the performance evaluation of month #1. The employees will be involved actively, asking them for feedback and suggestions on how to improve the sheets and how to reflect their actual work in the performance sheets better. The company's strategy, key objectives and the expectations of each function will be made clear to the employees. The involvement of the employees will also increase the chance of acceptance of the entire system amongst the employees. If attributes do not meet the requirements, the process for the particular attribute will restart from phase two. *Note: As BSC case studies such as Mooray et al. (1999) show, the first round of sheets will often be edited on towards a better round of sheets in the future.*

5. RESULTS: DEVELOPMENT OF THE PMS

The aim of the performance measurement system is to measure the performance of the employees, rather than the actual manufacturing, although the better the employee's performance, the better the expected quality of final products. The PMS should encourage a desired behavior amongst employees, a behavior that will contribute to the overall business strategy. The first four development steps of the model will be demonstrated in this chapter. These four steps combined will result in a first draft of performance evaluation sheets, these sheets will be used in the next chapters to demonstrate the impact of decision making.

PHASE 1: DERIVE KEY BUSINESS OBJECTIVES FROM STRATEGY

'01 Synergy' states the following:

> It is our mission to deliver defect-free software and services in a timely manner, to both internal and external customers. This will ensure high customer satisfaction. At '01 Synergy' we are totally committed to add value to our customers business by providing timely, cost effective & technological equivalents of planned obsolescence. We will comply with our Quality Management System and continually strive towards its improvement, as we believe quality is our key competitive differentiator.

'01 Synergy' believes the customer is the center of their business model. With multiple projects running at the same time, the company is interacting with multiple customers with diverse wishes at the same time. In order to succeed as a company it is essential to satisfy the customers by understanding their needs, and delivering work that answers their needs. This leads to the first key business objective; customer satisfaction.

The second key business objective that can be derived from the overall business strategy is 'quality assurance'. The company believes their quality of work is their key competitive differentiator. In order to maintain high quality standard '01 Synergy' requires a quality management system whereas the top management and supervisors can control the quality standards. High quality standards will also benefit the overall customer satisfaction.

'01 Synergy' also believes that in order to be able to achieve the required customer satisfaction and quality standards, a professional business culture is required. Professional behavior of the employees should be encouraged, in both the development labs where there is no direct contact with the client as in the business process outsourcing offices where employees may have direct contact with clients. This results in the third key business objective 'internal professionalism'.

Finally, in order to maintain the competitive quality edge and in order to be able to keep serving the customers the latest technologies in their fields, research and development should be used Employees should continue to improve their knowledge, especially in rapidly changing environments like the IT industry. Programmers should stay up-to-date with the latest developments in their field, and trainees should show considerable improvements. This results in the final key business objective 'personal development'.

'Table 7: BSC to business objective' compares the retrieved objectives with the four perspectives of the BSC. The financial perspective of the BSC does not result into a business objective in our

model. This PMS measures performance of *individual humans* rather than performance of *production units*. Financial values can be assigned to full production units, but problems may arise when trying to assign certain financial values to individuals, especially when considering the lack of project management and documentation at '01 Synergy'. This perspective may be reconsidered in future versions of the PMS.

TABLE 7: BSC TO BUSINESS OBJECTIVE

BSC Perspective	Business Objective
Customer	Customer Satisfaction / Quality Assurance
Learning and Growth	Personal Development
Internal Business Processes	Internal Professionalism / Quality Assurance
Financial	-

PHASE 2: CONVERT KEY BUSINESS OBJECTIVES INTO MEASURABLE ATTRIBUTES

During the next phases, the 'developers' sheet as found in Annex 5: Developer Sheet will be used as the example. Since the design process is the same for every function within '01 Synergy', solely the creation of the measures for developers will be explained in detail. This is also the reason why quality assurance will not be discussed in this chapter, since that key objective only includes measures for the functions at human resources and business process outsourcing. The other sheets can be requested at the author, an overview of the final set of found attributes linked to their business objectives is given in Annex 3: Strategy to attributes.

CUSTOMER SATISFACTION

Since most of the work at '01 Synergy' is project based, the attributes that define customer satisfaction are variable. The attributes may change per project in order to guarantee the satisfaction of the particular client. Customer satisfaction is derived in multiple measures, of which three are related to the developers, as shown in Table 8. The three measures will be averaged per month. This approach was chosen since the amount and complexity of the development work may vary highly per project, but it is expected that the amount of work and complexity of work will be about equal for each developer in a month.

1) *The amount of customer complaints/bugs after releasing their work.* Often it is not clear whether the reported problem should be qualified as a bug, i.e. a mistake on the developers end, or as a communication error with regards to the required end product. It may be, for example, that a customer qualifies something as a bug, but the developer qualifies it as a change in design rather than an error in the code. However, since customer satisfaction should be guaranteed, both cases will be considered and since the cut off between a bug and a design/communication error is hard to determine, they both will be combined in this measure.
2) *On time delivery*. Days of delay per assigned task. Average the total scores per task in a month to get the monthly scores. If a delay is in the upper limit of the scale, the score will be also at the upper end of the score and visa versa.

3) *Understanding of the project description*. The desired behavior of this measure is that developers quickly and accurately translate the needs of the client into the development of the project. The score will be determined by the amount of interactions per project. An interaction is defined as one email back and forth, however short follow up emails can be counted for the same interaction. If multiple projects start within a month, the monthly score will be the average of scores per project.

TABLE 8: DEVELOPERS ATTRIBUTES CUSTOMER SATISFACTION

Attribute	Measurement Formula
1) Customer Complaints and Bugs	Amount of required changes per month
2) On Time Delivery	Days of delay per assigned task, averaged per month
3) Understanding of Project	The amount of client's response interaction per project, between the project overview made by the developer and the actual start of the programming. Averaged per month. *Note: All contact between client and developer is per email.*

TABLE 9: SCORESHEET CUSTOMER SATISFACTION

Score	Customer Complaints and Bugs	On Time Delivery	Understanding of Project (score per project)
10	Less than 5	On time	One interaction
9 <= X < 10	Between 5 and 10	Delay <= Half a day	Two interactions
8 <= X < 9	Between 10 and 20	Half a day < Delay <= 1 day	Three interactions
7 <= X < 8	Between 20 and 30	1 day < Delay <= 2 days	Four interactions
6 <= X < 7	Between 30 and 40	2 day < Delay <= 3 days	Five interactions
X < 6	More than 40, 10 changes subtracts one grade	More than three days of delay. One day delay subtracts one grade	More than five interactions, with one interaction subtracting a grade

INTERNAL PROFESSIONALISM

The ability to maintain a professional internal culture is one of the key factors in creating an efficient organization. This business objective includes the set of common objectives; they apply to all employees.

TABLE 10: DEVELOPER INTERNAL PROFESSIONALISM ATTRIBUTES

Attribute	Measurement Formula
Attendance	(Time in office) / (Expected time in office, excluding excused leave) * 10
Punctuality	Daily Reports on time, deadlines met per month. Negative marks for submitting too late, tracking sheet maintained by HR, with the

	exception of the sheet for HR, which is maintained by the GM
Discipline / Distraction	Negative marks for personal calls in the office, Using PC for personal use, late comings etc. Tracking sheet maintained by HR, with the exception of the sheet for HR, which is maintained by the GM
Interpersonal Relations	General interpersonal behavior of people within the office. Number of interventions by HR. Measured per month
Group Performance	The average individual performance score of all members in a particular team for the measured month.

The score for attendance is determined by (Time in office) / (Expected time in office, excluding excused leave) * 10. The score of the attribute Group Performance is calculated by averaging the total individual scores of all team members in a particular team. For example if three developers score 7, 8 and 9 in a month, the group performance score will be the average of these three, and thus an eight. This was chosen to motivate each individual member in a team to not only improve their own performance, but also improve the performance of every single member in their team which should contribute to the overall team performance. The other attributes are scored as in the table below, however it should be noted that the below table is a guideline due to possible differences in importance of a remark. The scorer is therefore allowed to deviate from the table if required and if explained well. A heavy interpersonal incident, for example, may be valued much higher than a small argument and thus more marks can be deducted.

TABLE 11: SCORESHEET DISCIPLINE, INTERPERSONAL RELATIONS, PUNCTUALITY

Score	
10	No negative marks
9 <= X < 10	1 remark
8 <= X < 9	2 remarks
7 <= X < 8	3 remarks
6 <= X < 7	4 remarks
X < 6	More than 4 remarks, with one grade subtracted per remark

PERSONAL DEVELOPMENT

The key business objective personal development closely related with the learning & growth perspective of the BSC. Most of the measures derived from this objective will be assigned to the trainees. However, when possible future personal development plans for each employee will be created, additional attributes for different functions can be added. Currently developers are encouraged to increase their knowledge by self-training in order to stay up-to-date with the latest developments in their fields. The score for self-training is calculated as shown in the table below, with the hours without work being tracked by HR, and the hours of actual training done are tracked in the computer usage logs. In case of no free time, the attribute will be assigned a zero weight for the month, and not included in the overall score. Note that the developers themselves are not responsible for the scheduled training hours, and thus these scheduled training hours are not included in the metrics of their measure. The executives should schedule

free hours for training, so a measure for the executives can be 'Free available hours for training / total FTE'.

TABLE 12: DEVELOPER PERSONAL DEVELOPMENT ATTRIBUTES

Attribute	Measurement Formula
Self-Training Done	(Hours of self-training done at times there is no project work) / (recorded hours without work) * 10

PHASE 3: TESTING THE RELIABILITY AND VALIDITY OF THE ATTRIBUTES AND DOCUMENTATION OF THE ATTRIBUTES

At first the scorer of the company should be tested for reliability. Testing the reliability of *the single scorer* is done by selecting three developer scores for each month and calculating the $mean_{ij}$ of the scores of each employee i at all attributes related to key objective j for the particular month. The mean of a key objective j is determined by the average of all attributes related to the key objective, and the means are shown in Table 13. Thus the $Mean_{11}$ of month May is the mean of the attributes related to key objective 1 for employee 1 in May.

TABLE 13 SCORER RELIABILITY MEANS

Month	May	June	July
$Mean_{11}$	8.38	4.75	5.50
$Mean_{21}$	9.25	7.88	9.13
$Mean_{31}$	8.50	8.00	9.50
$Mean_{12}$	8.20	9.00	8.60
$Mean_{22}$	8.60	9.30	9.00
$Mean_{23}$	8.20	8.40	9.80
Mean of the Means	8.50	7.90	8.60

The average change in mean is below 10%, thus the scorer is found to be a reliable scorer.

Table 14 tests the reliability of the found attributes, whereas the typical error as measurement as a Coefficient of Variation (%) (CV) should be below 15.0% to pass the test.

TABLE 14 ATTRIBUTE RELIABILITY

Measure	Typical error as measurement as a CV (%)
Customer Complaints and Bugs Details	8.0
On Time Delivery	11.0
Understanding of Project	11.8
Self-Training Done	**45.3**
Attendance	3.4
Punctuality	12.2
Discipline	7.9
Interpersonal Relations	7.0

Self-training done does not show up consistent results and too high variation. It does not match the criteria of a standard error as a CV below 15%, and thus is determined to be unreliable and will be removed from the set of attributes until the measurement formula is optimized to ensure reliability.

When apply the framework of Neely *et al.* (2002) to the found attributes, the quality of the attributes can be tested. If any of the fields of the framework, particularly the formula and target ones, cannot be filled the quality of the measure should be questioned. The measure may encourage the wrong behavior then due to the attribute being vague and expectations unknown.

The framework applied to the measure 'Understanding of Project & Timely Response' of the developers is shown below. The other measures can be found in Annex 6: Testing the measures.

TABLE 15: UNDERSTANDING OF PROJECT & TIMELY RESPONSE

Understanding of Project & Timely response Details	
Title	Understanding of Project & Timely response
Purpose	To be able to quickly understand & translate the wishes of the client into the development of the project
Relates to	Business Objective 'Customer Satisfaction'
Target	Within one interaction
Formula	The number of client's interactions per project, between the project overview made by the developer and the actual start of the programming. *Note: All contact between client and developer is per email.*
Frequency of measurement	Per Project
Frequency of review	Monthly
Who measures?	GM
Source of data	Mail Contact with Client
Who acts on the data?	GM
What do they do?	Train the employees in possible culture differences between the employee and the general client base and improve logical reasoning.
Notes and comments	If the project size and complexity varies too much from the average project size, both determined by the GM, the GM may choose to lower or increase the amount of required interaction.

PHASE 4: INTERPRETATION OF THE ATTRIBUTES

The final step before the first draft of the performance measure sheets is determining weights using SMART's swing weights. The most important attribute will be assigned a value of 100, and the other attributes will be assigned a value compared to the most important attribute. If an

attribute is valued 50% less important than the most important attribute, the value will be 50. Finally the weight is determined by dividing the value of the attribute by the sum of the values of all attributes multiplied by 100. The final weights for developers is presented in the column on the far right, the other numbers are just used for calculation. The values are determined during valuation meetings with the involved managers (general manager & HR) and the team leader of the particular team, in this case the technical lead.

TABLE 16: APPLYING SWING WEIGHTS

Attribute	**Determine Weights**	
	Value	Norm
Customer Complaints and Bugs	100	35
On Time Delivery	100	35
Understanding of Project & Timely response	85	30
Total	**285**	**100**
Attendance	100	40
Punctuality	30	12
Discipline	30	12
Interpersonal Relations	30	12
Group Performance	60	24
Total	**250**	**100**

6. PERFORMANCE DECISION MAKING

When trying to improve the performance of the people within the company, the company can focus on one of the four found business objectives. An increase in either of them will lead to an increase in the performance of the company, ceteris paribus. However, the objectives also relate to other objectives. Furthermore, the measures linked to an objective may also relate to measures of other objectives. The purpose of this chapter is to give insights in these relations, and thus to show the dynamics of the system.

The causal relations between the key business objectives are shown in Figure 8: CLD Objectives. The figures should be read as 'The performance score of attribute X will result in a higher or lower performance score of attribute Y'. Quality Assurance, for example, benefits all other business objectives. A too high focus on customer satisfaction may however have a negative impact on the internal professionalism. People may ignore their administrative tasks and their professional behavior in order to catch a deadline which may eventually result in a drop in customer satisfaction.

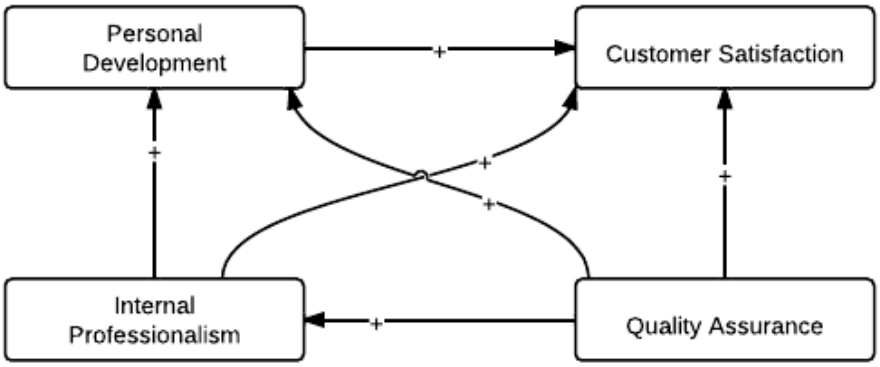

FIGURE 8: CLD OBJECTIVES

Customer Satisfaction and Internal Professionalism will be explained on a deeper level, again considering just the attributes of the developers. Firstly, having a better understanding of the project is likely to benefit both on time delivery and customer complaints. A too high focus on meeting the deadlines could result in a lower quality of work and thus a lower score on the attribute and visa versa.

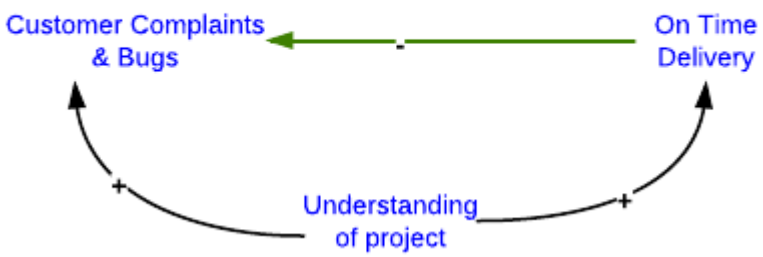

FIGURE 9: CLD CUSTOMER SATISFACTION

The causal relations for the attributes of Internal Professionalism are more complicated, and are shown in Figure 10. A higher attendance will benefit the interpersonal relations and the group performance. A higher score for interpersonal relations will also result in a better group performance. An increase of focus on punctuality and discipline, however, is expected to result in worse interpersonal relations.

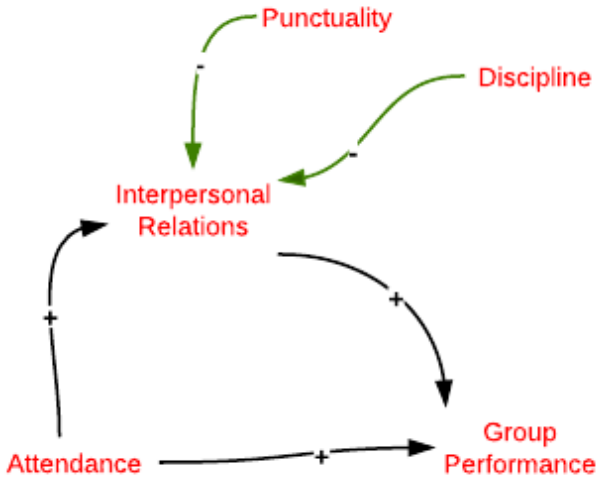

FIGURE 10 INTERNAL PROFESSIONALISM

6.1 DETERMINATION OF THE WEIGHTS

The final step of the simulation model is the determination of the weights. After adding weights to the causal relations and applying these to the overall calculation of the score, possible feature scenarios can be graphed. The weights can be determined partly by the earlier found swing weights, and partly by logical reasoning. It is noted that the accuracy of the predictions will greatly increase whenever these weights can be proven empirically.

The total score on a key business objective consists of two parts;

1. The direct influence of each attribute on the total score of the related key objective.
2. The interrelated growth coefficient, which represents the internal influence of each attribute on a different attribute belonging to the same key objective.

The first part can be constructed by using the weights that were determined in Phase 4 by converting the weights into a formula, whereas the coefficient of each attribute score is the same as the given weight in Phase 4 divided by 100.

So, for example, the total score of Customer Satisfaction is calculated by the following formula;

1. *Total Score Customer Satisfaction = 0.35 * Score (Complaints & Bugs) + 0.35 * Score(On Time Delivery + 0.3 * Score(Understanding of Project)*

The interrelated growth coefficient is based on the internal causality of each attribute as determined in the first part of this chapter. The best way to determine the values of this causality is to calculate the correlation scores between several attributes. However, due to the lack of data – the PMS was just introduced within this company and multiple months of scores

would be needed to determine reliable correlation scores – it was chosen to determine the weights by logical reasoning and expectations using Complaints & Bugs as an example.

As explained in Figure 9 on the previous page, the growth coefficient of the score at Complaints & Bugs is expected to be influenced by both changed in the Understanding of the Project and On time Delivery. A high focus on meeting the deadline will result in a lower quality of work and thus a high focus at On Time Delivery negatively influences a good result (low amount of complaints is a higher score) for Complaints & Bugs. Furthermore, a high Understanding of the Project would benefit the score for Complaints & Bugs since the quality of work is expected to increase whenever the understanding of the project increases. Following from this causality it was estimated to that 25% of the expected change in the score of On Time Delivery should be included in the final score on Complaints & Bugs (with a negative relation), and 10% of the expected change score at Understanding of the Project with a positive relation. This results in the following growth coefficient.

2. *Score (Complaints & Bugs)$_{n+1}$ = Score (Complaints & Bugs)$_n$ + [Score (Complaints & Bugs)$_n$ * Growth Coefficient(Complaints & Bugs)]*
    a. *Growth Coefficient(Complaints & Bugs) = Expected Δ(Complaints & Bugs) - 0.25*Expected Δ(On Time Delivery) + 0.1*Expected Δ(Understanding of Project))*
    b. *The growth coefficient depends on the chosen weights – or found correlation -, which in this case is -.25 for On Time Delivery, and .1 for Understanding of Project*

Illustrating this with hypothetical numbers whereas the expected changes in score is the variable entered when selected a certain policy ;

Expected Δ(Complaints & Bugs) = -3%
Expected Δ(On Time Delivery) = 5%
Expexted Δ(Understanding of Project) = 3%
Last month's Score (Complaints & Bugs)$_n$ = 5.00

Then;

Growth Coefficient = -0.03 – 0.25*0.05 + 0.1*0.03 = -0.0395 = -3.95%

*Total Expected Change in Score = 5.00 **-0.0395 = -0.1975

Score (Complaints & Bugs)$_{n+1}$ = 5.00 -0.1975*= 4.8025*

It should be noted that the maximum score at an attribute is never allowed to exceed 10, and thus this should be entered in the formula.

IF(*Score (Complaints & Bugs)$_{n+1}$ <= 10, Score (Complaints & Bugs)$_{n+1}$, 10)*

The excel sheet 'Developer Dynamics.xlsx' allows tuning future projections at different policies. One scenario will be discussed in here to show the functionality of the sheets.

6.2 SCENARIO ANALYSIS

In this hypothetical case our team of developers' performance below average in the past month with a score of 5 out of 10 at every attribute. The company has to pay a fine for every day of delay and new goals are determined with a high focus on meeting the deadlines. The executives determined that the performance of on time delivery should increase with 3% and

understanding of the project with 3%. For this, the executives expect that the amount of customer complaints & bugs will increase (and thus the score decrease). The model then shows the projects of the performance scores in the next twelve months based on the given weights. Despite a decrease of the score of complaints & bugs from 5.0 to 3.28, the total performance score on the key business objective Customer Satisfaction is expected to increase from 5.0 to 6.10 over a 12 months period, (see Excel sheet) as shown in Figure 11.

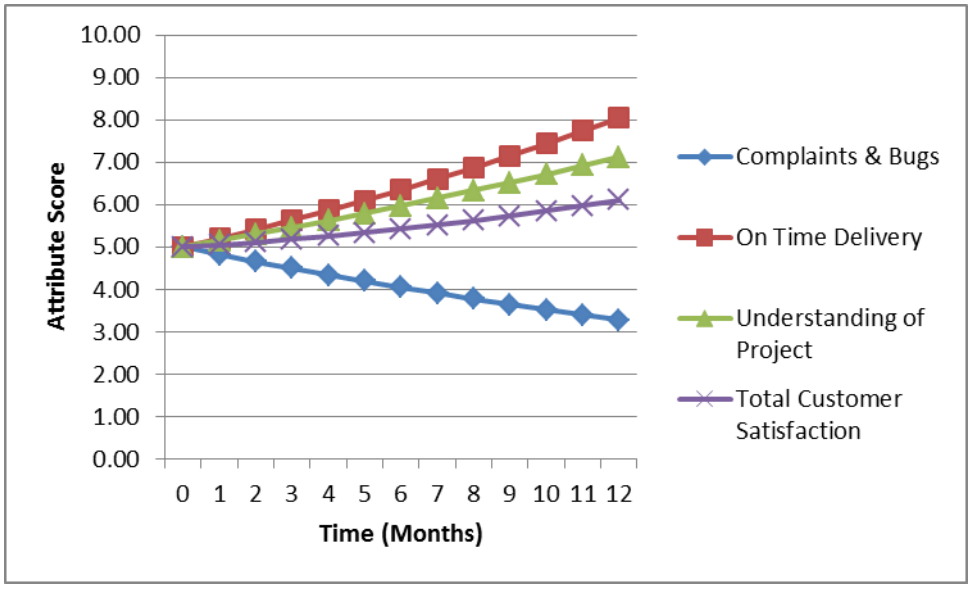

FIGURE 11: CUSTOMER SATISFACTION PROJECTIONS

At the same time, the company is unhappy with the discipline and the punctuality of the employees, and aims to increase the performance of punctuality with 3% per month, and discipline with 2% per month. They decide to lower the focus on interpersonal relations. The expected results from choosing this policy are shown in Figure 12. While punctuality and discipline are expected to increase due to the shift in focus, both group performance and interpersonal relations will greatly decrease. The total performance score on Internal Professionalism is expected to decrease over a 12 month period when choosing this policy.

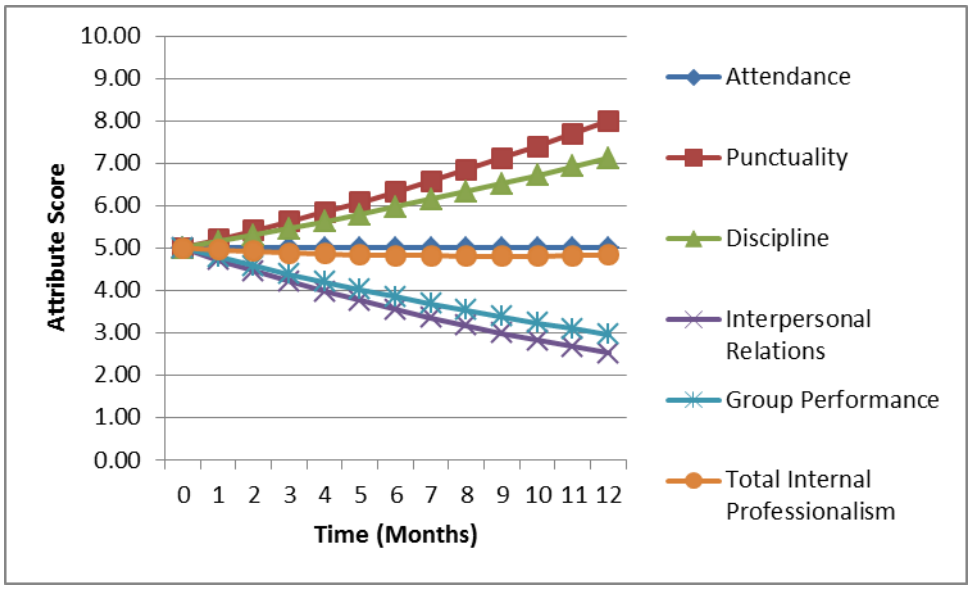

FIGURE 12 INTERNAL PROFESSIONALISM PROJECTIONS

As shown in the final Figure 12, these two policies - or expectations – combined are predicted to result in an increase from 5.42 to 6.06 within twelve months.

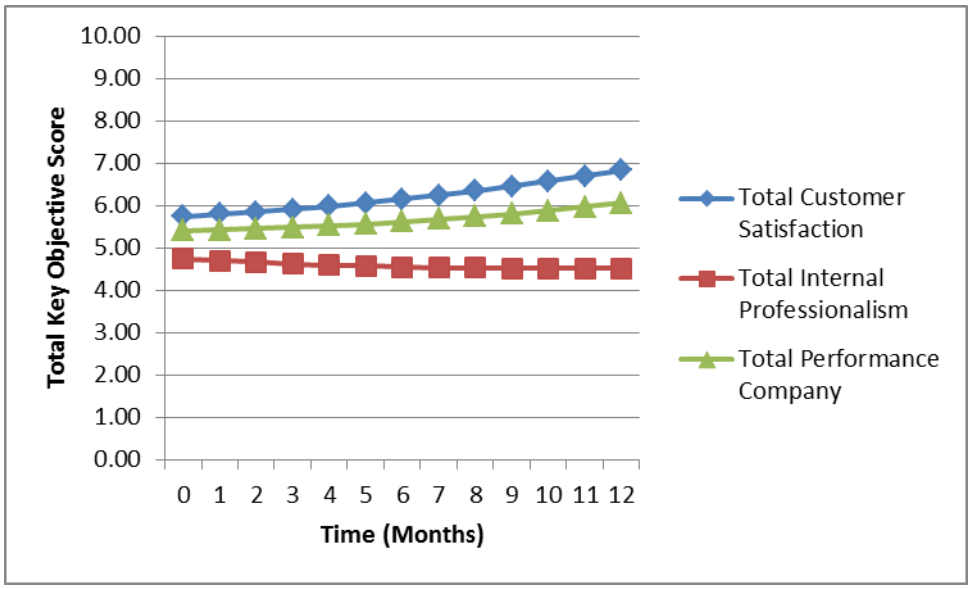

FIGURE 13: TOTAL PERFORMANCE PROJECTIONS

Given the model, when choosing the above stated hypothetical policies, both the punctuality and discipline of the employees within the company is expected to increase, however the total performance for Internal Professionalism is expected to decrease due to this policy. A higher focus for on time delivery, however, is expected to increase the overall customer satisfaction over a 12 month period, even though the amount of complaints/bugs is expected to increase. The executives should certainly revisit the chosen policy with a high focus on punctuality and discipline, since the impact of that decision is shown decrease the total internal professionalism.

7. EVALUATION OF THE SYSTEM

Hevner, A. R., S. T. March, et al. (2004) suggest several methods in order to evaluate the method of a designed system, as shown in Figure 14. In the past chapter the experimental evaluation method was already conducted. By experimenting with the model, simulations of certain policies were made. Next to the experimental evaluation method, the model will also be tested in practice, as shown in this chapter, in order to show the limitations of the model.

Table 2. Design Evaluation Methods	
1. Observational	Case Study: Study artifact in depth in business environment
	Field Study: Monitor use of artifact in multiple projects
2. Analytical	Static Analysis: Examine structure of artifact for static qualities (e.g., complexity)
	Architecture Analysis: Study fit of artifact into technical IS architecture
	Optimization: Demonstrate inherent optimal properties of artifact or provide optimality bounds on artifact behavior
	Dynamic Analysis: Study artifact in use for dynamic qualities (e.g., performance)
3. Experimental	Controlled Experiment: Study artifact in controlled environment for qualities (e.g., usability)
	Simulation – Execute artifact with artificial data
4. Testing	Functional (Black Box) Testing: Execute artifact interfaces to discover failures and identify defects
	Structural (White Box) Testing: Perform coverage testing of some metric (e.g., execution paths) in the artifact implementation
5. Descriptive	Informed Argument: Use information from the knowledge base (e.g., relevant research) to build a convincing argument for the artifact's utility
	Scenarios: Construct detailed scenarios around the artifact to demonstrate its utility

FIGURE 14: DESIGN EVALUATION METHODS [24]

A few problems appeared when testing the first set of performance measures in month one. First of all the sheets contained both common and function-specific measures. Due to a different amount of function-specific measures per function, the total weight of the common measures also differed per function as a result of applying the swing weights system. The weight for common measures should be the same for all functions, and thus it was decided to split the performance sheets in two parts, the common part and the function-specific part. In the set new performance sheets, employees will be able to get two separate weighted average grades, one for the set of function specific attributes and one for the set of common attributes. The grade for common attributes would count for 30 percent of the total grade, while the grade for the function specific attributes would count for 70 percent of the total. The common and function specific attributes were assigned two different swing weights sets, neglecting the influence of the amount of attributes to the weights of the common attributes so that all values on common attributes would be the same for every function.

Secondly, the general manager was unsatisfied with the weight for the common attribute Group Performance. Some teams were performing badly due to people denying their responsibilities and shifting the responsibilities to their teammates. This most certainly led to a lower customer satisfaction since deadlines were harder to meet, and the quality of work decreased. The general manager aimed to improve the employee's behavior in these teams and thus requested to

increase the weight of the attribute Group Performance. After discussing this suggestion in the review meeting, it was agreed on to double the value of Group Performance from 30 to 60.

Thirdly, both the general manager and the HR executive found it difficult to explain the employees that their performance should improve, even when they discussed the final grades per attribute with the employees in the monthly meeting, and that the general understanding amongst employees that their performance was excellent was wrong. This resulted in changing the method of the sheets a little. In the new system, employees would be asked one by one to fill in their own grades for the past month first. Then HR would fill in their grades for each employee and then the grades could be compared in the personal monthly review meetings. This way the difference in the employees' perception of their work and expectations of the management for the performance of the employees could easily be compared and explained. The other advantage of this system change is that it would increase the support and acceptance of the system amongst employees.

Fourthly the one of the employees' suggested attributes was included in the performance measurement sheets, the common attribute 'interpersonal relations'. Several employees valued the social abilities of their colleagues, and suggested it would increase the overall productivity if the overall interpersonal relations were excellent. When adding this as an attribute you may encourage employees to invest in the relations with their colleagues. While there would most certainly arise problems with setting targets and defining the measurement clearly, eventually the measure was still accepted and included in the performance measurement sheets.

Finally, the simulation model clearly showed some errors. Extreme cases cannot be executed with use of this model, and improvements should be found so that even extreme cases can be simulated. Whenever values like an expected decrease of for example 55% for a certain attribute, and an expected increase of f.e. 55% of another attribute of the same key objective within a month are simulated with help of the model, then predictions are unreliable. Inter causal relations should be improved in order to be able to simulate accurate and reliable predictions, even with extreme values. When these problems are overcome, and the method in finding the causal weights/values for the attributes is improved, the model would be able to accurately show the executives the impact of decision making, which now still is a shortcoming of the current system.

8. CONCLUSION & DISCUSSION

In industries with a high level of environmental changes, like the IT Outsourcing industry of 01 Synergy, performance measurement systems allow organizations to align its daily operations to the overall company strategy. When performance measurement systems are integrated successfully the organizations may change in the direction of a learning organization that is able to adapt to the high variability of its environment. Even though several authors acknowledge the value of such systems, most performance measurement frameworks are static, whereas the performance measurement systems should by dynamic in order to be effective. It is of high importance to have a performance measurement system in place that should be easily modified in order to maintain relevant measures. Furthermore, the system should not just be able to evaluate the past by the generated data on the selected attributes, but the system should also be able to predict the impact of decision making. Most of the current performance measurement systems lack both these aspects.

When applying System Dynamics to static performance measurement systems such as the Balanced Scorecard, the static systems can be converted to dynamic performance measurement systems. System Dynamics increase insight in the behavior of the system by showing causal relations and by simulating the impact of certain chosen policies. Whenever System Dynamics are applied correctly, the executives are also able to show the impacts of changing certain variables and thus increase modifiability and adaptability of the system to maintain its relevance. After creating a causal loop diagram by determining causality between attributes, a model can be created based on the causal relations. Weights can be assigned to demonstrate the value and causality between the attributes, which allows the creation of a simulation model. After finishing the model, possible policies can be tested, and the expected impact of each policy can be graphed. Using such models will assist the decision makers in choosing the correct policy.

Introducing the performance measurement system is just the foundation of continuous improvement and transferring into a learning organization. Future personal development plans should be created in order to guarantee the learning & growth of the organization. Continuous improvement and review of the PMS is needed in order to maintain the relevance of the attributes. By optimizing both the weights for the attributes itself, and the causal relation weights between the attributes, the system's accuracy is expected to increase considerably. Future development of the system is required, research on possibly applying empirical correlation tests should be done to improve the weights. Finally, the system is not able to handle extreme numbers such as an expected decrease of 55% in on time delivery per month due to the causal relations. Future additions to the system should be able to address this issue so that the error margin can be decreased and the accuracy of the predictions can be increased.

9. BIBLIOGRAPHY

1. Kaplan, R. and D. Norton, *The Strategy Focused Organization*2001, Boston: Harvard Business School Press.

2. Bourne, M., et al., *Implementing performance measurement systems:a literature review.* Int. J. Business Performance Management, Vol. 5, No. 1,, 2003: p. 1-24.

3. Akkermans, H.A. and v. Oorschot, K E, *A case study of Balanced Scorecard development using System Dynamics.* Journal of the Operational Research Society 56, 2005: p. 931-941.

4. Mayer, R.C. and J.H. Davis, *The Effect of the Performance Appraisal System on Trust for Management:A Field Quasi-Experiment.* Journal of Applied Psychology, 1999: p. 123-136.

5. Neely, A., *The performance measurement revolution: why now and what next?* International Journal of Operations & Production Management, 1999. **19**(2): p. 205-228.

6. Hudson, M., A. Smart, and M. Bourne, *Theory and practice in SME performance measurement systems.* International Journal of Operations & Production Management,, 2001: p. 1096-1115.

7. Mooraj, S., D. Oyon, and D. Hostettler, *The Balanced Scorecard: a Necessary Good or an Unnecessary Evil?* European Management Journal, 1999: p. 481–491.

8. Santos, S.P., V. Belton, and S. Howick, *Adding Value to Performance Measurement by Using System Dynamics and Multicriteria Analysis*, 2001, Strathclyde Business School: Glasgow.

9. Gardiner, P.C. and A. Ford, *Which policy run is best and who say so?* TIMS Studies in Management Sciences 14, 1980: p. 241-257.

10. Forrester, J.W., *Industrial Dynamics*, 1961, MIT Press: Cambridge, Massachusetts.

11. Sterman, J.D., *System Dynamics: Systems thinking and modeling for a complex world*, 2003, Massachusetts Institute of Technology Engineering Systems Division: Massachusetts.

12. Mustajoki, J., R. Hamalainen, and A. Salo, *Decision Support by Interval SMART/SWING.* Decision Sciences, 2005: p. 317-339.

13. Goodwin and Wright, *Decision Analysis for Management Judgement*, 1998, John Wiley & Sons: Chichester.

14. Kaplan, R.S. and D.P. Norton, *The Balanced Scorecard - Measures that drive performance*, 1992, Harvard Business Review: Cambridge.

15. Bourne, M., et al., *The success and failure of performance measurement initiatives.* International Journal of Operations & Production Management, 2002. **22**(11): p. 1288-1310.

16. Edward G. Carmines, R.A.Z., *Reliability and validity assessment*, 1979, Sage Publications: California.

17. Neely, A., et al., *Designing performance measures: a structured approach.* International Journal of Operations & Production 17(11), 1997: p. 1131-1152.

18. Kennerley, M. and A. Neely, *A framework of the factors affecting the evolution of performance measurement systems.* International Journal of Operations & Production Management, 2002. **22**(11): p. 1222-1245.

19. Peffers, K., et al., *A Design Science Research Methodology for Information System Research.* Journal of Management Information Systems, 2008: p. 45-77.

20. Kaplan, R.S. and D.P. Norton, *Linking the Balanced Scorecard to Strategy.* California Management Review, 1996: p. 53-79.

21. Kaplan, R.S. and D.P. Norton, *Transforming the Balanced Scorecard from Performance Measurement to Strategic Management: Part 1.* Accounting Horizons, 2001: p. 87-104.

22. Bourne, M., et al., *Designing, implementing and updating performance measurement systems.* International Journal of Operations & Production Management, 2000: p. 754-771.

23. Lebas, J.M., *Performance measurement and performance management.* International Journal of Production Economics, 1995: p. 23-35.

24. Hevner, A.R., et al., *Design science in information systems research.* Mis Quarterly, 2004: p. 75-105.

10. ANNEX

ANNEX 1: TERMINOLOGY & DEFINITIONS

BSC – The Balanced Scorecard

MCDA – Multi Criteria Decision Analysis

SD – System Dynamics

CLD – Causal Loop Diagram

PMS – Performance Measurement System

BPO – Business Process Outsourcing

DM – Decision Maker

Productivity – 'The total power of '01 Synergy' to complete the current projects according to the client's wishes.'

ANNEX 2: PERFORMANCE MEASURE RECORD SHEET

Details
Title
Purpose
Relates to
Target
Formula
Frequency of measurement
Frequency of review
Who measures?
Source of data
Who owns the measure?
What do they do?
Who acts on the data?
What do they do?
Notes and comments

FIGURE 14: PERFORMANCE MEASURE RECORD SHEET (NEELY. ET AL (1997))

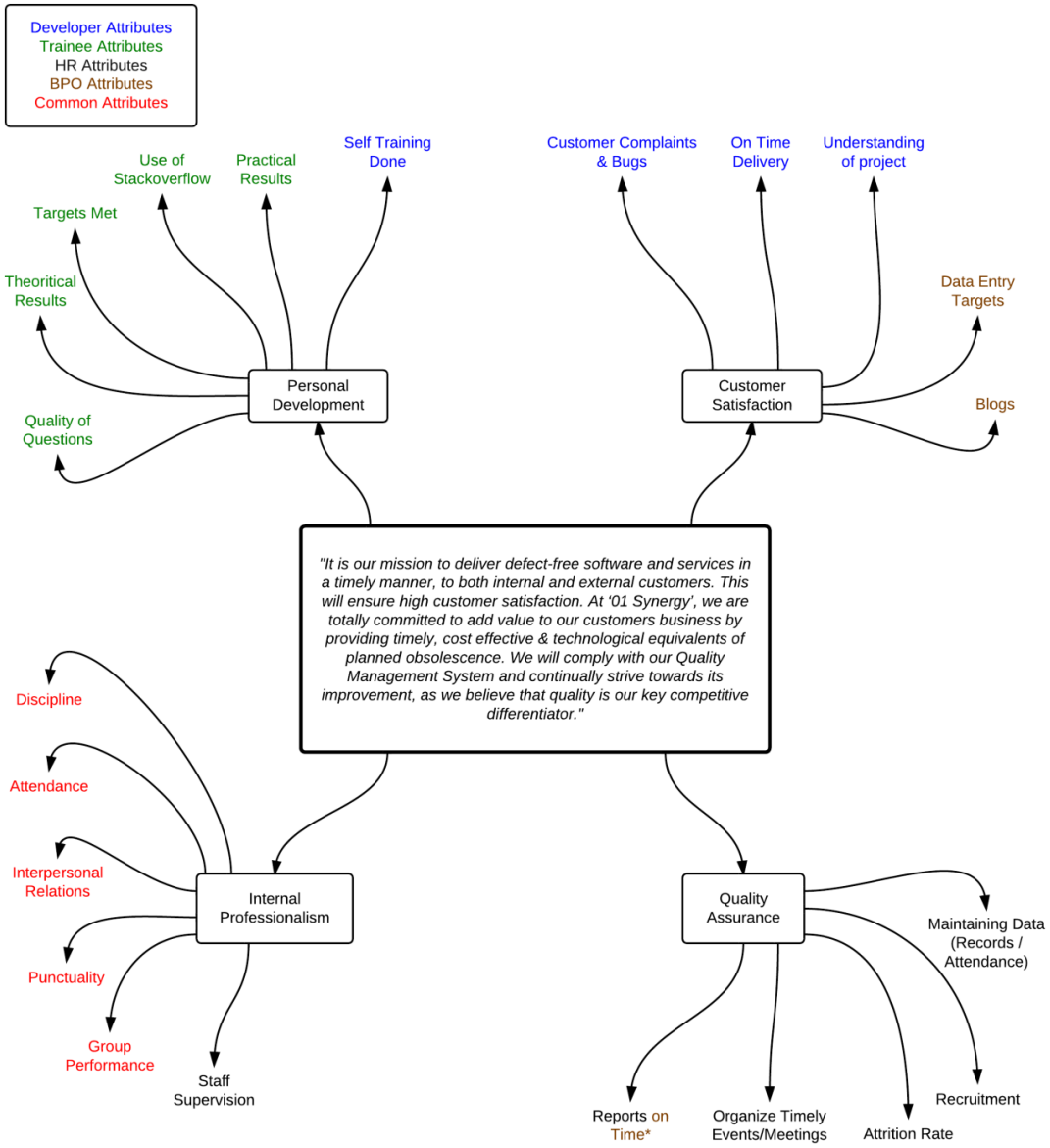

FIGURE 15: OVERVIEW OF ATTRIBUTES AND OBJECTIVES

ANNEX 4: EXPLANATION OF ADDITIONAL LOOPS

TABLE 17: N2. 'NEW EMPLOYEES LOOP'

N2. 'New Employees Loop'.	
Loop Explanation	This loop is connected via "A". At a high need of resources, the work pressure will be high, possibly increasing the hiring rate of new employees, increasing the amount of new staff, finally leading to a lower need of resources. (Note the maybe hard readable '– (minus)' from 'New Staff' to '# Needed Resources'.)
Current Situation	The work pressure for the development teams is high in 01 Synergy. Synergy is actively recruiting for developers, however, recruitment for these fairly new technologies appears to be challenging.
Perfect Situation	A lower need of resources means more continuity within the company and a lower need for recruitment.
Difference Between Current and Perfect Situation	Ineffective recruitment. Insufficient resources.
Involved Actors	HR Executive, General Manager
Notes	A standard, although not cheap, solution for correcting a lack of resources. It must be noted, however, that hiring new staff might not bring the desired increase in productivity, as further explained in 'P3 – Pressure Turnover Quality Loop'. It takes time to find the employees and to integrate them within the company culture.

TABLE 18: P2. 'SATISFACTION PRODUCTIVITY LOOP'

P2. 'Satisfaction Productivity Loop'	
Loop Explanation	This loop shows a different vicious circle around productivity. A high productivity leads to a lower need for resources, leading to a lower work pressure, leading to higher employee satisfactions, again leading to a higher productivity. This process can be turned around as well; a low productivity will lead to an even lower productivity.
Current Situation	Productivity is low, work pressure is high. Employees are not satisfied with the high work pressure due to much overwork and extra work at weekend days in order to complete projects.
Perfect Situation	Enough resources so that expensive and demoralizing overwork is redundant.

Difference Between Current and Perfect Situation	Too much overwork. Not enough resources.
Involved Actors	HR Executive, General Manager, Employees.
Notes	N/A

TABLE 19: P3. PRESSURE TURNOVER LOOP

P3. 'Pressure Turnover Loop'	
Loop Explanation	The aim of this loop is to show the value of experienced people. If turnover rates increase due to a lower employee satisfaction, the amount of experienced people will decrease, leading to a lower quality of resources eventually again leading to a higher turnover read. This also works the other way around. The lower the turnover rate, the higher the average level of experiences amongst the employees, the higher the quality of resources, the better the productivity, thus less workload, higher satisfaction and an even lower turnover rate.
Current Situation	There is a lack of experienced people. The trainees cannot develop into full employees, and experienced developers are hard to recruit.
Perfect Situation	Maintain the experienced developers, recruit more experienced people. Lower work pressure. Allow trainees to develop into full-time programmers.
Difference Between Current and Perfect Situation	Not enough experience. Inefficient training plans. Ineffective recruitment.
Involved Actors	HR Executive, General Manager.
Notes	N/A

ANNEX 5: DEVELOPER SHEETS

See attached Developer Evaluation Sheet.xlsx

See attached Developer Dyanmics.xlsx

Customer Complaints and Bugs Details	
Title	Customer Complaints and Bugs
Purpose	To improve the quality of work offered to the customer
Relates to	Business Objective 'Customer Satisfaction'
Target	No changes required
Formula	Amount of required changes after delivery of the project per month
Frequency of measurement	Monthly
Frequency of review	Monthly
Who measures?	General Manager
Source of data	Client Contact
Who acts on the data?	General Manager
What do they do?	Improve the work of the developers
Notes and comments	

On Time Delivery Details	
Title	On Time Delivery
Purpose	To meet the set deadlines
Relates to	Business Objective 'Customer Satisfaction'
Target	No delays
Formula	Days of delay per assigned task
Frequency of measurement	Monthly
Frequency of review	Monthly
Who measures?	General Manager
Source of data	Project Schedule
Who acts on the data?	General Manager
What do they do?	Encourage better planning and encourage developers to deliver the work on time
Notes and comments	

Attendance Details

Title	Attendance
Purpose	To encourage people to be on time and work full hours
Relates to	Business Objective 'Professionalism'
Target	100% Attendance
Formula	Time in office vs. expected time in office, excluding excused leave.
Frequency of measurement	Monthly
Frequency of review	Monthly
Who measures?	HR Executive
Source of data	Attendance Register
Who acts on the data?	HR Executive
What do they do?	Identify reasons for absenteeism and try to improve daily attendance
Notes and comments	

Punctuality Details	
Title	Punctuality
Purpose	To document your work well
Relates to	Business Objective 'Professionalism'
Target	Below 5 marks
Formula	Daily Reports on time, deadlines met per month. Negative marks for being too late.
Frequency of measurement	Monthly
Frequency of review	Monthly
Who measures?	HR Executive
Source of data	Punctuality Sheet
Who acts on the data?	HR Executive
What do they do?	
Notes and comments	

Discipline Details	
Title	Discipline
Purpose	To keep the office an environment that allows people to concentrate
Relates to	Business Objective 'Professionalism'
Target	Below 5 tally marks

Formula	Negative Marks for personal calls in the office, Using PC for personal use, late comings etc. Measured per month.
Frequency of measurement	Monthly
Frequency of review	Monthly
Who measures?	HR Executive
Source of data	Discipline Sheet
Who acts on the data?	HR Executive
What do they do?	
Notes and comments	HR Executive uses tally marks

Interpersonal Relations Details	
Title	Interpersonal Relations
Purpose	To improve communication and business spirit at the work floor
Relates to	Business Objective 'Professionalism'
Target	Full Marks
Formula	General interpersonal behaviour of people within the office. Negative marks when anything unusual occurs with intervention of HR. Measured per month.
Frequency of measurement	Monthly
Frequency of review	Monthly
Who measures?	HR Executive
Source of data	Attendance Register
Who acts on the data?	HR Executive
What do they do?	Identify interpersonal problems and find ways how to solve it.
Notes and comments	

Group Performance Details	
Title	Group Performance
Purpose	To encourage team members to increase each other's and their own effectiveness
Relates to	Business Objective 'Professionalism'
Target	Total Group score above 9.0
Formula	The average performance score of all members in team for the measured month.
Frequency of measurement	Monthly

Frequency of review	Monthly
Who measures?	HR Executive
Source of data	Performance Sheets
Who acts on the data?	HR Executive
What do they do?	Identify team's problems and increase team harmony
Notes and comments	