Table of Contents

Go Back

Developing a Balanced Scorecard with System Dynamics

Henk Akkermans and Kim van Oorschot
Eindhoven University of Technology
Department of Technology Management
Paviljoen F.18
P.O. Box 513
5600 MB Eindhoven
The Netherlands
tel.nr. +31 40 2472230, fax nr. +31 40 2464596
h.a.akkermans@tm.tue.nl , k.e.v.oorschot@tm.tue.nl

Abstract

The Balanced Scorecard (BSC) is a popular concept for performance measurement, because it focuses attention of management on just a few measures and bridges different functional areas (both financial and non-financial measures are included in the BSC). But, the BSC has also received some criticism. In this paper five limitations of the use of the BSC are discussed: BSCs focus on unidirectional causality, are unable to distinguish delays between actions and their impact on performance, have a dearth of validation capabilities, integrate insufficiently strategy with operational measures and suffer from internal biases. We propose a system dynamics approach to develop a BSC in order to overcome these limitations. We present a case study from the insurance sector where this approach is applied. The results suggest that developing a BSC with system dynamics is a promising approach to supplement existing BSC frameworks.

Keywords: balanced scorecard, system dynamics simulation, insurance industry, case study

1. Introduction

Performance management, and in particular the version of it known as balanced scorecard (BSC), has become a topic of considerable interest in both the business world (Siegele 2002) and in academia (Kaplan and Norton 2001). The term and concept of the balanced scorecard originates from the field of Management Accounting, where control on the basis of purely financial criteria was found to be inadequate (Kaplan and Norton 1992). From there on it has quickly become popular in other areas of management research, such as organisation studies (Dinesh and Palmer 1998), operations management (Neely et al. 1995, Bourne et al. 2000) and information systems (Martinsons et al. 1999).

The reasons for this sudden rise to prevalence are obvious: On the one hand, there is the appeal of simplicity: no longer do managers have to work their way through piles of statistics, but they can monitor some 5-10 key indicators instead. On the other hand, there is the strength of interdisciplinarity. In the past, all the relevant inputs from different functional areas had to be translated into financial data. Some functions, such as marketing, were considerably more suited to do this than others, such as R&D or operations (Hill 1989). But, regardless of how successful this translation was, it remained a conversation in a foreign language. With a BSC, managers have now a more acceptable common language to discuss issues in.

Upon reflection it becomes clear that underlying the BSC concept there are some strong, albeit usually implicit, systemic notions. Only when one looks at the organisation from a systems approach (Churchman 1969) one starts to understand why anyone would think that monitoring a dozen or so performance indicators can be sufficient to manage something as complex as an organisation. The implicit belief here has to be that, of the thousands of observable variables and their interrelations, only certain very specific causal linkages or loops will be dominant in determining overall system behaviour. Another relevant, albeit implicit, systems notion is that of interconnectedness (Ackoff 1981): the belief that all different aspects and functions of the organisation are interrelated and that one cannot improve one area, or the whole for that matter, without influencing all other areas as well. Hence, the call for an interdisciplinary approach.

Despite or perhaps in response to its popularity, the BSC concept is not without criticism. In this paper we will argue that these criticisms can again be traced back to an often implicit systems perspective on organisations. Broadly speaking, these criticisms can be seen as the other side of the coin for the before-mentioned advantages of the BSC. Yes, from a systemic perspective one may expect that only a few process indicators can point at the key leverage points of the system (Forrester 1992, Sterman 2000). But, how can one be sure that the dozen or so selected are indeed the right ones? Should there be more? Or less? And, do they all work in the same direction or might they counteract each other? Moreover, if they are the right variables, what would be the right values to target for? And within what time frame should these be achieved? If everything is interconnected, then where should the system boundaries be placed? From a theoretical perspective, these are not trivial questions. Nor are they, in practice, trivial to managers actually implementing a BSC, as we will see in this paper.

In this paper we discuss five shortcomings to the currently common way of developing and using a BSC. This discussion is based upon a review of the BSC literature. We also present a BSC development approach to overcome these weaknesses, one that is based upon system dynamics modelling.

Moreover, we describe the application of this approach in a case study. This case concerned the development of a BSC for a company in the insurance industry, where the management team of a business unit formulated a BSC to facilitate implementation of its recently-formulated business strategy. As our analysis of this case will show, most of the shortcomings mentioned in the literature could be overcome with our approach.

The remainder of this paper is structured as follows. The upcoming section contains a brief literature overview of the main contribution of the BSC and the shortcomings noted to current theory and practice. Then we discuss how system dynamics modelling could overcome these shortcomings. Section 4 describes the case setting and Section 5 the structure of the simulation model that was developed. We then present in Section 6 the findings from our model analysis. These illustrate how, in this particular case, most of the noted shortcomings were actually overcome. Research limitations, opportunities for further research and practice development are all discussed in Section 7. Our main conclusions wrap up this paper.

2. Advantages and limitations of current BSC theory and practice

The Business scorecard is a performance measurement system introduced by Kaplan and Norton (1992). According to these authors, a BSC addresses shortcomings of traditional performance measurement systems that relied solely on financial measures. To overcome this, Kaplan and Norton (1992) introduced three additional measurement categories that cover non-financial aspects. The result is a scorecard that translates the vision and strategy of a business unit into objectives and measures in four different areas:

1. *The financial perspective*: how the company wishes to be viewed by its shareholders;
2. *The customer perspective*: how the company wishes to be viewed by its customers;
3. *The internal business process perspective*: in which processes the company must be adept in order to satisfy it shareholders and customers;
4. *The organisational learning and growth perspective*: which changes and improvements the company must achieve to implement its vision.

The "balance" of the scorecard is reflected by the balance between lagging (outcome measures) and leading (performance drivers) indicators, and between financial and non-financial measures (De Haas 2000).

A BSC links measures of process performance, or key performance indicators (KPI’s), together in a causal chain that leads through all four perspectives: measures of organisational learning and growth influence measures of internal business processes, which, in turn, act upon measures of the customer perspective, which ultimately drives financial measures (Nørreklit 2000). Causal chains or causal diagrams provide a medium by which people can externalise mental models and assumptions and enrich these by sharing them (Wolstenholme 1999). In fact, according to Neely (1998), one of the hidden strengths of a balanced measurement framework, in particular of the BSC, is that it forces management teams to explore the beliefs and assumptions, which underpin their strategy.

The BSC concept originates from the U.S.. There it has been applied successfully across many industries and within the public sector. It has also been delivered to an international audience, on a multi-disciplinary front (Hepworth 1998). For example, Malmi (2001) found that the logic of the BSC was appealing to many companies in Finland. Wisniewski and Dickson (2001) describe its application to a police force in Scotland.

Partly, the success of the scorecard can be explained by the right timing and marketing. Managers were frustrated with traditional measurement systems at the time when the BSC was promoted in a series of articles in the Harvard Business Review (Kaplan and Norton 1992). But, furthermore, the BSC appears simple. It suggests that only a few well-balanced numbers are required to measure the performance of a company. And, the BSC can serve as a bridge between different fields (financial and non-financial fields). These two advantages are further discussed below.

Advantage #1: Checking just a few numbers

The first advantage of the BSC is that just a few numbers or performance indicators need to be checked. Generally, the BSC prescribes that only three to five measures should be developed for each of the four perspectives mentioned in the previous paragraph (Neely 1998). The BSC is therefore more narrow than the Management by Objectives (MbO) approach, which is seen as a predecessor of the BSC (Dinesh and Palmer 1998). MbO requires eight areas in business where performance objectives must be set. These areas are: market standing, innovation, productivity, physical and financial resources, profitability, manager performance and development, worker performance and attitude, public responsibility (Pugh and Hickson 1989). As a consequence, MbO is more complex than the BSC and this is seen as an important reason why successful implementation of MbO has been disappointing (Dinesh and Palmer 1998).

This is not surprising from the perspective of bounded rationality. It has long been known that rationality of human decision making is best described as “bounded”, rather than optimal (Simon 1957; Cyert and March 1963). This is caused by cognitive limitations to knowledge, abilities to process information, and limits of time. Indeed, Miller (1956) has shown that for most humans short-term working memory is limited too "7±2" chunks of information. In many organisations usually too many measures exist to be digestible for humans. The BSC forces managers to focus on only four business areas and within each area only on the most important performance indicators. From a cognitive point of view, this is a distinct advantage.

Advantage #2: Bridging the gap between different fields

The second advantage of the BSC is that it serves as a bridge between different fields. Both financial and non-financial measures are included in the scorecard. Also researchers from different management fields have examined the concept. The management accountancy aspect of the BSC has been considered by, for example, Newing (1994) and Nørreklit, (2000). Also in the operations management field the BSC is well-known (Neely et al. 1995; Bourne, et al. 2000; Hafeez et al. 2002, Lohman et al. 2002) From a strategy perspective, the BSC has been described by for example Mooraj, et al. (1999), Hudson, et al. (2001) and Kaplan and Norton (2001). Furthermore, the concept has been used for the strategic management of information systems (Martinsons, et al. 1999). This interest in and the successful use of the BSC by researchers and managers from different fields indicates that it is possible to combine performance measures related to different aspects of a company into only one scorecard.

Limitations

Next to the well-published successes of BSC a number of inherent weaknesses have been reported in the BSC literature. Interestingly, the advantages of the BSC mentioned in the previous paragraph can also be interpreted as disadvantages:

1. Unidirectional causality too simplistic. The use of causal-loops alone is seen as problematic because these loops do not capture the notion of strategic factors accumulating and depleting. Moreover, there is little basis in a causal-loop map for estimating the scale or speed of change of key items (Warren and Langley 1999). Nørreklit (2000) even questions the existence of a causal relationship between the areas of measurement in the BSC. Instead of a causal relationship, this author believes that the relationship is more one of interdependence, or bi-directional causality.

2. Does not separate cause and effect in time. Nørreklit (2000) points at the problems arising from the fact that the time dimension is not part of the BSC, because in some cause-and-effect relationships a time lag exists between cause and effect. This is not shown by the BSC since it measures cause and effect at the same time. Platts and Kim (2002) agree that simply looking at different measures simultaneously is not enough. The linkeages between them must also be understood.

3. No mechanisms for validation. The BSC concept provides no mechanism for maintaining the relevance of defined measures (Hudson et al. 2001). Neely et al. (1995) also found that the problem for managers is usually not identifying what could be measured, but reducing the list of possible measures to a manageable (and relevant) set. Thus, the advantage of checking just a few numbers may become a disadvantage when not the right numbers are selected for the BSC. Furthermore, the advantage of bridging the gap between different fields may become a disadvantage when performance indicators of different fields counteract or thwart each other.

4. Insufficient links between strategy and operations. Mooray et al. (1999) state that the BSC fails to identify performance measurement as a two-way process. It focuses primarily on top-down performance measurement. Hudson et al. (2001) also acknowledge this and write that BSCs have a lack of integration between the top-level, strategic scorecard, and operational-level measures.

5. Too internally focused. A BSC may be too narrowly defined. Mooray et al. (1999) discuss that the BSC does not consider the extended value chain in which employee and supplier contributions are highlighted. Neely et al. (1995) argue that the BSC is not able to answer one of the most fundamental questions for managers: what are the competitors doing? Thus, the advantages of checking just a few numbers related to different fields may become a disadvantage when important fields are overlooked.

3. System dynamics as a method to overcome BSC limitations

In this article, we advocate the use of system dynamics (SD) as a method to overcome the limitations to current BSC theory and usage that are mentioned in the literature. In particular, we suggest a two-stage process of SD model building for BSC development:

- *Stage 1*: Elicit mental models from management of perceived interrelationships using causal loop diagrams (CLDs). Generate a BSC on the basis of the discussions this mapping process is leading to.
- *Stage 2:* Translate causal loop diagrams into a quantified simulation model using key company data. Test BSC on the basis of this simulation model with managers and discuss implications for mental models and BSC.

This two-stage approach to system dynamics model building is, in its generic format, a "normal" and accepted way of conducting system dynamics interventions in organisations (c.f. Senge 1990, Lane 1992, Winch 1993, Vennix 1996, Wolstenholme 1999, Sterman 2000, Akkermans 2001). In this paper, we aim to advance its application to the process of developing a BSC. Further on, we will present a case of an insurance company business unit where such a BSC development was conducted successfully using system dynamics. In the remainder of this section, we will focus on the question to what extent this approach can be instrumental in overcoming the before-mentioned limitations to current BSC development and usage.

1. Feedback loops rather than unidirectional causality. In the real world, causality is rarely unidirectional. A not only influences B, but B, over time, also indirectly influences A via, for instance, C and D. They are connected in a feedback loop, which is a cornerstone of system dynamics thinking and modelling. Unfortunately, , it has long been noted that when managers "spontaneously describe their environment, they do not include (…) feedback cycles" (Axelrod 1976, p.238). Fortunately, Axelrod also notes that the same policy makers "have no trouble accepting the separate beliefs that make up a feedback cycle when each link is given explicit attention." (ibid, p.238). In this area, the OR literature has long recognised that "system dynamics provides an aid to sensitising policy makers to the feedback cycles and their implications by constructing influence diagrams" (Hall and Menzies 1983, p.53). [Influence diagrams are nowadays usually called causal loop diagrams (Sterman 2000).]

2. Explicit separation of cause and effect in time. As essential as feedback loops to the system dynamics worldview is the notion of delays between cause and effects. Delays are generated by accumulations or "levels" as they are called in SD terminology and it is delays that generate instability in dynamic systems. Therefore, in SD cause and effect are by definition separated in time. All but too often, measures intended to improve a certain situation first lead to deterioration in performance and this effect is routinely captured in system dynamics models.

3. Mechanisms for rigorous validation. A long-debated issue within the system dynamics community is if the first stage of modelling has sufficient value on its own, without subsequent translation of the qualitative insights derived into a quantified simulation model that can be more rigorously tested, validated and investigated (e.g. Wolstenholme and Coyle 1983, Wolstenholme 1999, Coyle 2000, Sterman 2000, Homer and Oliva 2001). But, in the context of BSC, the relevant observation to be made is that, in system dynamics modelling, it is common practice to quantify and thereby validate qualitative assumptions about strategy and policy. For it remains well-proven that people in general are not at all proficient in inferring correct dynamic behaviour from qualitative structures (Morecroft 1983, Sterman 1989, Paich and Sterman 1993). Moreover, even if they are able of thinking through the dynamic result of a separate policy of striving for a particular value of a particular KPI, when these separate policies come together in execution, the intentionally rational separate policies may well give rise to unexpected and undesirable outcomes (Morecroft 1985, Forrester 1992).

4. Linking strategy with operations. System dynamics models are commonly characterised as especially successful in capturing strategic issues, rather than tactical or operational ones (e.g., Akkermans and Bertrand 1997). But this does not mean that SD models contain no links to operational processes and performance indicators. On the contrary: constructing an SD model starts with identifying the main operational flows in an organisation and the main stages in those flows: the flow of customer orders, of goods or services, of employees, of money, of capital goods etc (Forrester 1961, Richmond 1994). Specific elements of strategy, or policies, are then always translated into specific values for specific parameters that drive the rates in those flows. "Operational thinking" (Richmond 1994) is at the core of system dynamics modelling and can be highly instrumental in bridging gaps between strategic KPI's on the one hand, and operational processes on the other.

5. Broadening focus by challenging system boundaries. If many current BSCs can be criticised for being too internally focused, then system dynamics models are often prone to comments that their scope is too wide. Perhaps combining the BSC concept and SD modelling can help in achieving a better balance here. At any rate, it is good modelling practice in system dynamics modelling to challenge the nature of every exogenous variable: is it really exogenous or, in one way or another, indirectly influenced by the variables that are endogenous to the model being developed? (Sterman 2000). This process of "challenging the clouds", as it has been called by

Barry Richmond, the developer of the leading SD-software package Ithink, is pertinent to the system dynamics approach (Richmond 1994).

4. The case study

The organisation investigated was part of a large insurance company. It was set up not so much as a normal business unit, but as a separate legal entity, to emphasise its independence from the parent's other businesses. The reason for this was the nature of its services: it provided legal aid to clients who were entitled to such services on the basis of their insurance policies.

In the past few years, this organisation had gone through a period of considerable upheaval. First there had been several changes in top management. Then it was faced with a major increase in demand for its services, as a result of changes in the market and different insurance sales policies with its parent. In response, staffing was increased significantly, after a long period of little to no growth. Recently, the organisation had undergone a major restructuring, shifting from a regional structure to structuring along different areas of judicial expertise, along lines of service one might say. Finally, the management team was almost completely new, i.e. less than a year in their current jobs. All in all, the time seemed right for a serious reorientation on key goals for the future.

This was the background against which our involvement with this company should be situated. We, the authors, formed part of a small group of external consultants who facilitated the development of a balanced scorecard for this organisation by the management team (MT). This development was set up in the two stages as described in the previous section.

During the first stage, preparatory interviews were conducted with MT members, the results of which were discussed in a half-day workshop where the group engaged in a number of causal loop diagramming exercises. The findings from this workshop were distilled in a so-called workbook (Vennix 1996), which the MT members filled in and sent back. These results were again discussed in a full-day workshop, where a first version of a BSC was created and agreed upon, which is shown in Table 1 without explicit targets for reasons of confidentiality.

Table 1. KPI's of the first version of the BSC

Key Performance Indicators	
output per employee	% of small and easy cases
throughput time per case	colleagues for colleagues
customer satisfaction	outsourcing of cases
employee satisfaction	number of successful projects
employee turnover rates	working at home
training on the job/coaching	

However, agreement in the management team went further: there was also agreement on the approach forward. Especially relevant in the context of the current paper is that the team felt pleased with their first BSC, but at the same time was uncertain about its validity. Were these really the right indicators? Had they been complete? And would they all work towards the same goal? To what extent would they really be instrumental in achieving the formulated company mission? On the other hand, could the list of KPI's be shortened? The fewer dials to watch the better, after all. To meet this uncertainty the authors proposed to develop a system dynamics model that would address all these understandable questions.

In the second phase of the project the two authors developed a quantified simulation model for this company. They did so on the basis of the causal loop diagrams developed in the first workshops and the BSC that had been determined. These were sufficient to develop a first skeleton of an SD model. This skeleton was filled with key company data, which were delivered by two managers from the MT. These two were more closely involved than the others in the subsequent development of the model, critiquing intermediate versions and providing valuable feedback. One of these, the internal project leader for the BSC development, performed what Richardson and Andersen have called a *gatekeeper role*: "a person within the client group who carries internal responsibilities for the project (…), helps frame the problem, (…) works with the modelling support team to structure the sessions and participates as a member of the group" (Richardson and Andersen 1995, p. 115).

The contents of the findings from this exercise are presented in the next section. Here is suffices to state that these findings have been discussed with the management team, have been challenged by them, in some cases have been mitigated but nevertheless have broadly been accepted. At the time of writing, these findings are being used to guide the implementation of the BSC approach for the organisation as a whole and well as for the various sub-units involved.

5. Model structure

As was described in the previous section, the second phase of the project was the development of the quantified simulation model of the insurance company. In this section the structure of the model is described. Because of reasons of confidentiality and size, we are not able to describe the model in detail. The quantitative relationships between variables are therefore not given in this paper. However, the full model (with fictitious numbers) is available from the authors upon request.

The simulation model is based on the causal loop diagrams that were developed by the MT members in the first phase of the project and the simulation model reflects the *production process* of the company, that is the processing of cases. Furthermore, also the *production capacity* is included in the model, that is the employees of this company. The loops that have been drawn by the MT members are feedback loops that connect the production process with the production capacity, or, in other words that connect the processing of cases with the employees. The model structure is drawn in Figure 1. In the next paragraphs the contents of the model are explained.

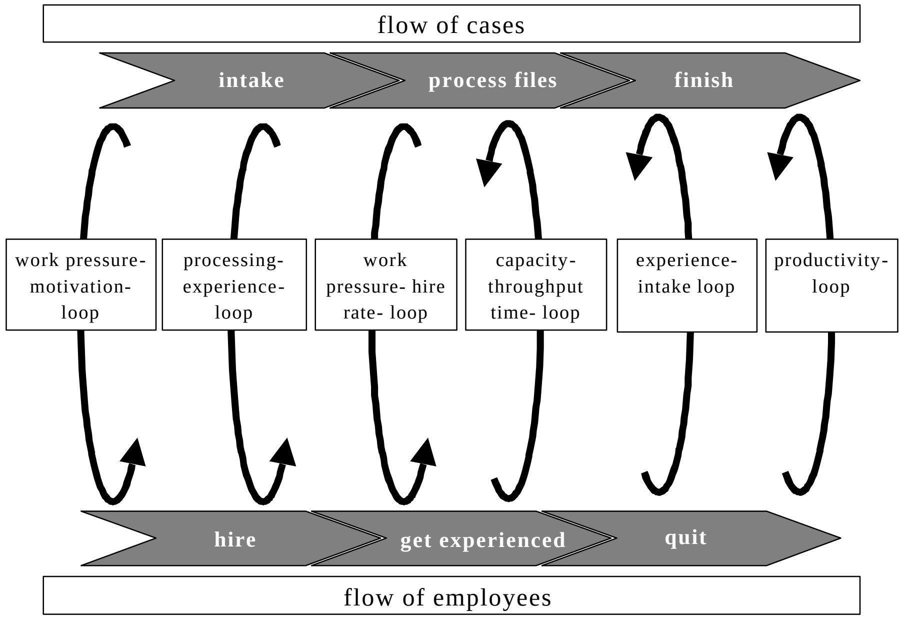

Figure 1. Overall model structure

Flow of cases

The legal aid that is provided by the insurance company can only be given after a so-called intake process. In this process it is decided whether the request for legal aid of a customer can be given by the company. When the intake is accepted, a specific customer file (a "case" in legal terms) is made and this case is allocated to an employee for further processing. Occasionally, when employees are too busy, a case can be outsourced. When the intake is rejected, the request of the customer is discarded without further processing taking place. This flow of cases is represented in stocks-and-flows notation (Sterman 2000) in Figure 2.

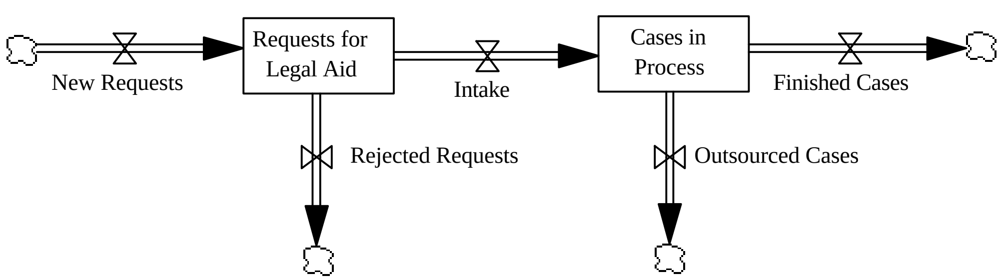

Figure 2. Stocks-and-flows diagram for cases

Flow of employees

Cases are processed by employees. The population of employees is divided into new and experienced employees. This distinction is necessary because both the turnover rates as well as the productivity differ for new and experienced employees. New employees become experienced after a certain assimilation time. In this company, the average assimilation time was three years. So, from the point of view of expertise, the productivity of experienced employees must be higher than that of new employees. However, experienced employees are required to train new employees. In doing so, experienced employees will 'lose' time also, thereby decreasing their productivity. The way in which productivity is determined in the simulation model will be described in the next paragraph. First, in Figure 3 the stocks-and-flows-diagram for the flow of employees is given.

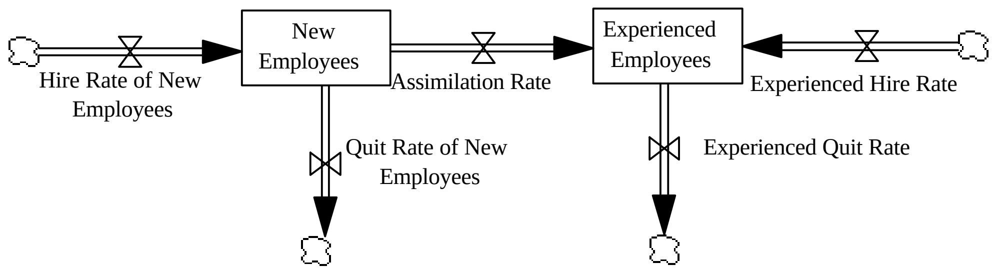

Figure 3. Stocks-and-flows diagram for employees

The productivity loop

As was shown in Figure 1, different loops connect the stocks and flows of cases with the stocks and flows of employees. In this paragraph we discuss the productivity loop in some more detail, since this is the most complex loop. In the latter part of this section, we will describe the other five loops in more general terms.

The connection between cases and employees through the productivity loop can be described as follows: the higher the experience of employees, the higher their productivity, and in turn, the higher the productivity the more cases employees can process, consequently the more cases are processed by employees, the more experienced they get, etcetera. Learning curve theory provides models to relate experience of employees with productivity. Here, we will use the model given by Sterman 2000, p. 507):

$$\text{Productivity} = \text{Reference Productivity} * \left(\frac{\text{Average Experience}}{\text{Reference Experience}} \right)^{c}$$

Average Experience of either the experienced or the new employees is the total experience (expressed in working years) divided by the number of experienced or new employees. Experience can be gained by processing cases, but experience can also be lost. People forget relevant knowledge and new developments in the insurance sector may cause experience to become obsolete. This is expressed in the model by an *experience decay rate* (in our model this rate is 10% per year). Furthermore, experience is also lost when employees leave the company. Therefore, the Average Experience can both increase and decrease over time. *Reference Productivity* is the productivity attained at the *Reference Experience* level. For example, in the simulation model the Reference Experience is about 7 working years for experienced employees and 0.2 working years for new employees. The Reference Productivity is about 200 cases per year per experienced employee and about 125 cases per year per new employee. The exponent c in the computation of the productivity determines the strength of the learning curve and is equal to

$c = \ln(1 + f_p) / \ln(2)$

where f_p is the fractional change in productivity per doubling of effective experience (Sterman 2000). Moreover, it is also included in our simulation model that effective experience can only be gained when employees are actually processing cases. This means that the more time employees spend on for example courses or training each other, the shorter the time that remains for processing cases and so the smaller the amount of experience that can be gained. Thus, attending courses and training colleagues may have a detrimental effect on productivity. On the other hand, the MT members also stated that attending courses and training colleagues may have a beneficial effect on motivation or job satisfaction. Therefore, the positive effect that motivation has on productivity is also included in our model.

Loops connecting cases and employees

Besides the productivity loop that was described above, five other loops that connect the stocks and flows of cases with those of employees are worth discussing.

Work pressure-Motivation loop. The number of cases that have to be processed per employee determine the work pressure that he or she perceives. From cognitive psychological literature it is known that when work pressure becomes too low or too high, people are less motivated to do their work and consequently their productivity is lower compared to highly motivated people (Yerkes and Dodson 1908, Fisher 1986). The higher the productivity, the more cases are processed, which in turn has a positive effect on the perceived work pressure.

Processing-Experience loop. This loop was discussed in some more detail before. It implies that employees can gain experience when they are processing cases. The more cases that employees process, the more experienced they will get, and in turn the faster they will be in processing cases in the future (higher experience will lead to higher productivity).

Work pressure-Hiring Rate loop. When the number of cases that have to be processed per employee is increasing, the work pressure increases too. When the work pressure becomes too high, the pressure to hire new employees also increases. This pressure to hire influences the hiring rate of new employees and, after a certain time delay, hiring new employees will decrease the work pressure of employees.

Capacity-Throughput Time loop. The higher the capacity that is available (thus excluding time spent on courses, training, holidays, etc.) for processing cases, the shorter will be the average throughput time per case. In turn, a short throughput time leads to more cases that are processed, which will cause experience of employees to increase. Higher experience will have a positive effect on productivity and a high productivity has a positive effect on the available capacity for processing cases.

Experience-Intake loop. In drawing the causal loop diagrams the MT members noted that when the Intake process is done by experienced employees, the quality of this Intake process is improved. Experienced employees know the strengths and preferences of colleagues and therefore they can ensure that the right case is allocated to the right employee (for example, a case concerning damage to a bicycle will be in capable hands when the employee is an avid cyclist). The better the fit between the content of the case and the interests of the employee, the higher his or her motivation and as a consequence, the higher the productivity. The higher the productivity the faster the case can be processed, leading in turn to more experience.

Together with the productivity loop described before, these loops connect the stocks and flows of cases and employees and make the simulation model a coherent model. In the next section a selection of the analyses performed with the model is discussed.

6. Model analysis

In the previous sections it was suggested that the use of system dynamics would be instrumental in overcome the mentioned limitations to current BSC development. In this paragraph the system dynamics model we introduced in Section 5 is analysed. For each of the five limitations that were described in Section 2 we show to what extent the use of a system dynamic model nullifies these limitations.

1. *Feedback loops rather than unidirectional causality*

It was discussed that causality is rarely unidirectional. As an example, we discussed six feedback loops. In Figure 4 the work pressure – motivation loop is shown. In this figure work pressure increases because the number of cases that needs to be processed grows. As a consequence of the high work pressure, employee satisfaction (or motivation) decreases rapidly. A low employee satisfaction also has a negative effect on the productivity (measured in C/(Y*E) = Cases per Year per Employee), that in turn also decreases. However, because of the increasing work pressure, more employees are hired (see Figure 5) and after a few years, this results in a positive effect on the work pressure. Only when work pressure has decreased enormously, a slight increase in employee satisfaction is shown.

It takes even longer for the productivity of employees to increase too. New employees have a low productivity and because new employees must be trained by experienced employees, the productivity of the latter is also negatively influenced on the short-term. When productivity rises, more cases can be processed and the work pressure can decrease even more.

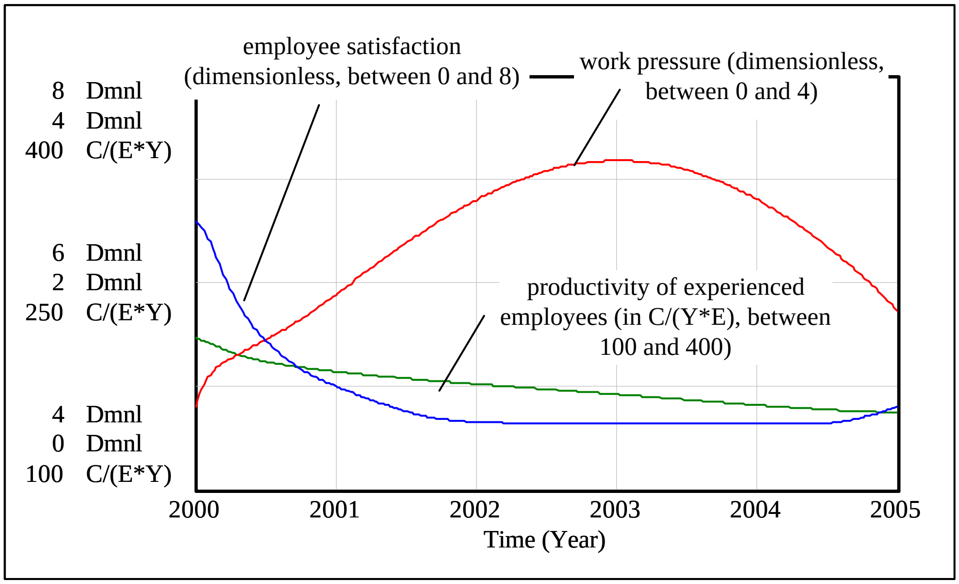

Figure 4: Work pressure – Motivation interactions over time

2. Explicit separation of cause and effect in time

In Figure 4 it was shown that work pressure increased because the number of cases that needed to be processed increased. This increasing number of cases also has an effect on the throughput time of cases, which is shown below in Figure 5. To decrease throughput time and work pressure, more employees are hired. However, it takes a certain time before new employees become productive and thus before the effect of the hiring of new employees becomes visible. For example, in 2000 the total number of employees is 207. In 2003 this number is 330. Between 2000 and 2003 the throughput time and the work pressure (see Figure 4) are still increasing. When time delays are not taken into account, it might be suggested that the hiring of new employees has no effect on the throughput time and work pressure at all. On the long term, on the contrary, hiring new employees does have an effect.

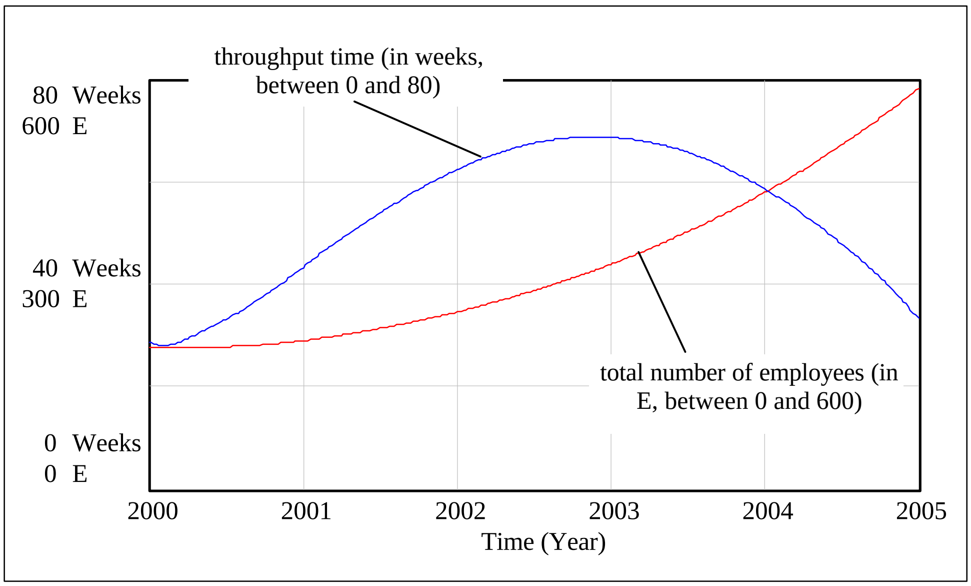

Figure 5. Throughput time and total number of employees

3. Mechanisms for rigorous validation

In Table 1 the KPI's that were selected by the MT members of the insurance company were shown. With the system dynamics model we tried to validate these KPI's. The results of this validation are shown in Table 2.

Table 2. Results of Validation of KPI's

Key Performance Indicator	Result of Validation Test
output per employee	valid
throughput time per case	valid
customer satisfaction	valid
employee satisfaction	valid
employee turnover rates	valid
training on the job/coaching	valid
% of small and easy cases	redundant
colleagues for colleagues	redundant

outsourcing of cases	redundant
number of successful projects	potentially contra productive
working at home	potentially contra productive

With an example we will explain how we validated the KPI's. “Outsourcing of cases” was believed to be an important mechanism to alleviate work pressure. The system dynamics model showed that indeed, on the short term, outsourcing has a positive effect on work pressure, as shown in Figure 6. In 2002, almost 15% of all cases are outsourced and indeed between 2001 and 2003 the work pressure is lower than in the case without outsourcing. However, after 2003, the work pressure is higher in the outsourcing-case. When cases are outsourced, employees are not able to gain experience with those cases, and on the long term this lack of experience has a negative effect on their productivity and consequently, their work pressure will be negatively influenced. This example shows that the average effect of outsourcing is not as positive as estimated. It was proposed to eliminate this KPI from the BSC.

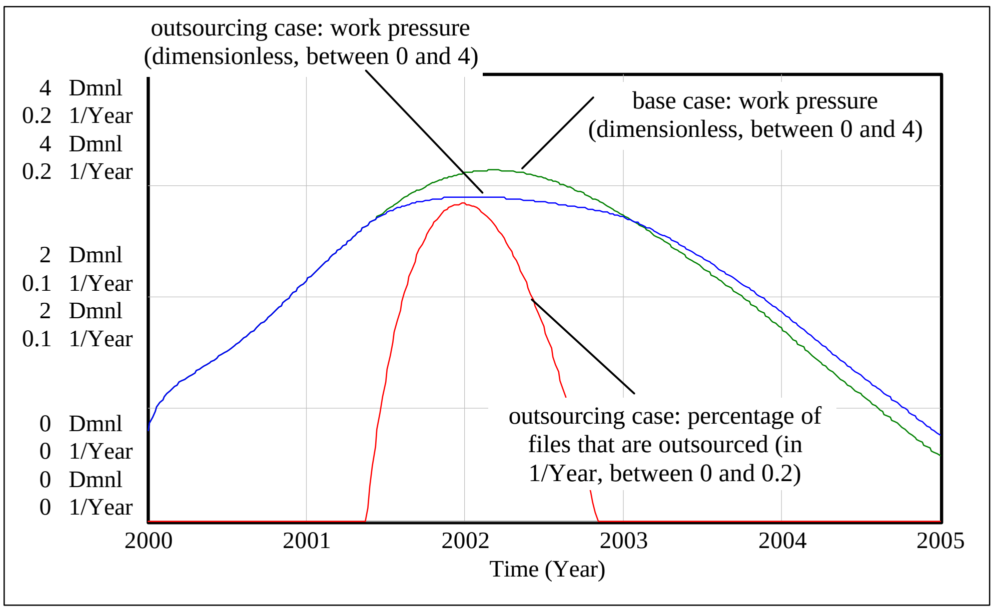

Figure 6. Effect of outsourcing on work pressure.

4. Linking strategy with operations

With the system dynamics model not only strategic measures can be tested, also operational measures get a chance. Increasing employee satisfaction, for example, is on of the strategic KPI's that are modelled. But to actually increase this satisfaction operational measures are required. The operational measure that was tested was the so-called *experienced intake* (see also the experienced-intake loop in Section 5). Experienced employees know the strengths and preferences of employees. When more experienced employees are involved in the intake of cases (in which cases are allocated among employees), the fit between the content of cases and the interests of the employees responsible for the case is better. The better the fit, the higher the employee satisfaction. This operational measure was tested. In the base case 50% of the intake of cases was done by experienced employees (and consequently 50% was done by new employees). In the experienced intake-case 75% of the intake was done by experienced employees, see Figure 7.

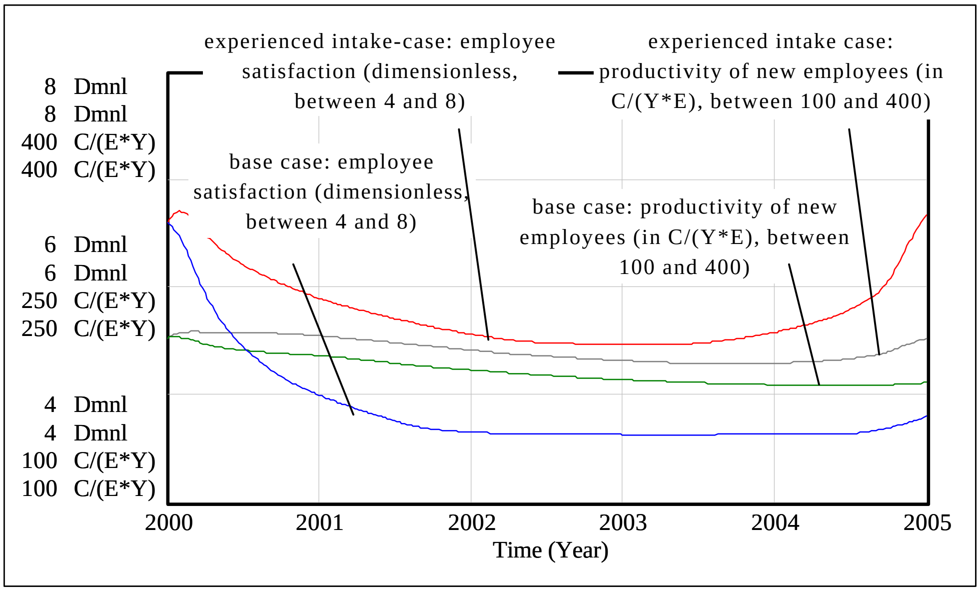

Figure 7. The effect of experienced intake

Figure 7 shows that the *experienced intake* has indeed a positive effect on employee satisfaction and consequently also on the productivity of employees.

5. Challenging system boundaries

The MT members of the insurance company made a Business Unit Plan that prescribed the number of employees that had to be hired in the next years (base case). To challenge this BU plan, it was tested what the result of a different hiring strategy would be. In the *reactive hiring case* the work pressure has a direct effect on the desired number of employees and thus on the number of new employees that are hired. The results are shown in Figure 8. (In this figure the *experienced intake* is 75%.)

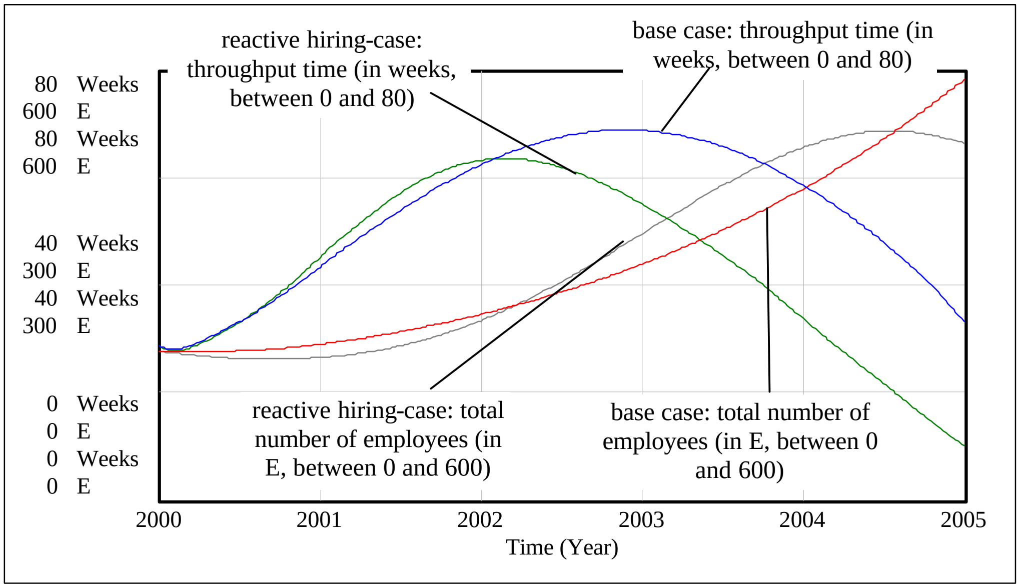

Figure 8. The effect of reactive hiring

Figure 8 shows that the difference between the total number of employees in both cases is zero during 2004. However, in the reactive hiring-case the total number of employees increases faster in 2002 and 2003, which has a positive effect on throughput times. Because throughput times are so low in 2005, the total number of employees stabilises, while in the base case this number is still increasing.

In this section five examples were shown of the use of the system dynamics model to develop the BSC. The next section discusses to what extent the five limitations of the BSC, which were described in Section 2, are actually nullified by the use of system dynamics.

7. Discussion

In this section we reflect on this study's findings. The previous section has emphasised the merits of our SD approach to BSC development. In this section we look more at the inherent shortcomings to this approach and possibilities and impossibilities for overcoming these in later studies.

Modelling of "mental maps", not of the "real world"

It is important to note that the process that we are suggesting for BSC development focuses on making explicit the mental maps of the individual managers, on sharing them, challenging their internal consistency and aligning them. What this approach does *not* attempt to do is to model the "real world", independent of what the managers perception of this real world is. The philosophical dimension of this distinction we will not solve in this article, as this goes back all the way to Plato's cave. There still remain two camps of academics: those who insist that all models are social constructions of reality and those who believe that there are at the very least significant elements of objectivity in all social system models.

The practical side can be easier resolved. Developing a rigorous model of real-world business processes through direct observation is a laudable but fairly time-consuming process. The approach that we have presented here is intended to supplement a strategic decision-making process. So, the fair comparison to be made is not between the model that one develops through the process we have outlined and some theoretical “optimal” model. Rather, one should set a BSC development approach with SD against the conventional approach of developing BSCs.

A focus on top management, not on other stakeholders

The approach described in the case study has strongly focused on the senior management team. Perhaps not surprisingly, at the time of writing the BSC that has been developed at the overall company level is still being translated into BSCs for the underlying units, and the concept of using it for personal scorecards is still in its infancy. The management team itself may have granted sanction to their scorecard, but the rest of the organisation still has to follow. Whether this is simply a matter of time, time will tell. Nevertheless, there seems to be a clear opportunity for improving this deployment of the BSC with the various mid-level management teams through engaging them in a similar process of group model-building as their senior management has gone through.

This points at two shortcomings of conventional BSC development approaches that we have not explicitly addressed in the preceding section, i.e. the absence of links between strategy and operations and their overly internal focus. Pursuing similar levels of insights and buy-in with mid and lower level employees by engaging them in similar processes of SD-enabled BSC development could remedy the first shortcoming. Involving other stakeholders such as suppliers, customers and employees to be incorporated in the BSC could overcome the second. Here it should be noted that, in general, SD-enabled group model-building processes have a good track-record of facilitating effective communication regarding complex issues between groups of diverse backgrounds (Vennix 1996, Akkermans 2001).

Links with other methodical approaches

System dynamics is a clear candidate to help overcome shortcomings of current BSC practice, but obviously not the only one. In response to the growing dissatisfaction with the shortcomings mentioned in our literature review, several different approaches are being developed at this time. One example is the work by Platts and Kim (2002). In general, any approach that starts from the existing BSC frameworks but adds to that effective group communication processes, visualisation techniques for causal linkages and opportunities for quantification and consistency-checking, will mean a considerable improvement for BSC practice and theory, we believe.

Opportunities for follow-up research

Our goal in this research was to demonstrate the benefits of using system dynamics in addressing shortcomings in current balanced scorecard development. For this we used our experiences from a single case study. Needless to say, the validity of findings from a single case is limited by definition. Therefore, this case should be seen as a first attempt to generate new theory regarding BSC implementation, rather than as a test of existing theories (Yin 1989, Eisenhardt 1989).

Obviously, follow-up research is required to refine and/or refute the propositions we are laying down in the current article. This could be either additional case studies using the same approach or comparisons of different BSC development methods. Survey research covering a larger population could perhaps shed more light upon the extent in which the shortcomings mentioned in the literature are also experienced in practice and under what circumstances these shortcomings are more or less severe. All this should happen relatively quickly if the BSC concept is not to become accused of faddism in the coming few years as a result of many unsuccessful implementations due to improper methodical support.

8. Conclusion

The balanced scorecard concept is an important one, both theoretically and practically. In this paper, we have seen that the main reasons for its wide appeal may well stem from its implicit systemic nature: its focus on a limited set of key performance indicators and its ability to align functional perspectives and objectives within the organisation. We have also seen that current BSC practice still suffers from serious shortcomings, which can be seen as resulting from inadequately dealing with this systemic nature. From our literature review appear five of such shortcomings: their focus on unidirectional causality, their inability to distinguish delays between actions and their impact on performance, their dearth of validation capabilities, their insufficient integration of strategy with operational measures and their internal bias.

From a systemic perspective, the improvements needed to take away these shortcomings are obvious: BSCs ask for both quantitative modelling and close involvement of more stakeholders than just the senior management team in the modelling process. Our choice for system dynamics as the modelling approach to do so is less idiosyncratic than it may seem at first. From its origins, SD has positioned itself halfway between strategy and policies and their operationalisation. Forrester (1961) talks already about the policy level, Morecroft (1984) about strategy support models (1984) and “microworlds for policymakers” (1992).

System dynamics remains, from the palette of systems interventions available, the technique that, in terms of quantitative modelling, was designed to “boldly go where no one has gone before” in areas where reliable data and theoretical models are lacking but nevertheless the need for simulation, for scenario analysis, is clearly apparent (Flood and Jackson 1991, Pidd 1996). The group model-building techniques aimed at group learning and consensus that have been developed in this field can be highly instrumental in broadening ownership and deployment of BSCs beyond senior management. Seen from this perspective, further synergy between SD and BSC looks like a win-win scenario for all stakeholders involved…

References

Ackoff, R.L. 1981. Creating the corporate future. Plan or be planned for. Wiley, Chichester.

Akkermans, H.A., Bertrand, W.J.M. 1997. On the Usability of Quantitative Modelling in Operations Strategy Decision Making. International Journal of Operations & Production Management 17(10): 953-966.

Akkermans, H.A., and Vennix, J.A.M. 1997. Clients’ Opinions on Group Model-Building: An Exploratory Study. System Dynamics Review 13(1): 3-31.

Akkermans, H.A. 2001. Renga: A systems approach to facilitating inter-organisational network development. System Dynamics Review 17(3): 179-194.

Axelrod, R. 1976. Structure of decision. Princeton University Press, Princeton NJ.

Bourne, M., Mills, J., Wilcox, M., Neely, A. and Platts, K. 2000. Designing, implementing and updating performance measurement systems. International Journal of Operations and Production Management 20: 754-771.

Coyle R.G. 2000. Qualitative and quantitative modelling in system dynamics: some research questions. System Dynamics Review 16(3): 225-244.

Cyert, R. and March, J. 1963. A Behavioral Theory of the Firm. Prentice-Hall: Englewood Cliffs.

Dinesh, D. and Palmer, E. 1998. Management by objectives and the Balanced Scorecard: will Rome fall again? Management Decision 36: 363-369.

Eisenhardt, K.M. 1989. Building Theories from Case Study Research. Academy of Management Review, 14 (4), 532-550.

Fisher, S. 1986. Stress and Strategy. Lawrence Erlbaum Associates, London.

Flood, R.L., Jackson, M.C. 1991. Creative problem solving. Total Systems Intervention. Wiley, Chichester.

Forrester, J.W. 1961. Industrial dynamics. MIT Press, Cambridge.

Forrester, J.W. 1992. Policies, decisions and information sources for modelling. European Journal of Operational Research 59(1): 42-63.

Haas, M. de. 2000. Strategic Dialogue: In Search of Goal Coherence. Thesis Eindhoven University of Technology: Eindhoven.

Hafeez, K., Zhang, Y. and Malak, N. 2002. Determining key capabilities of a firm using analytic hierarchy process. International Journal of Production Economics 76: 39-51.

Hall, R.I., Menzies W.B. 1983. A corporate system model of a sports club: Using simulation as an aid to policy making in a crisis. Management Science 29(1): 52-64.

Hepworth, P. 1998. Weighing it up – a literature review for the balanced scorecard. Journal of Management Development 17: 559-563.

Hill, T. 1989. Manufacturing strategy. The strategic management of the manufacturing function. Macmillan, London.

Homer, J. Oliva, R. 2001. Maps and models in system dynamics: a response to Coyle. System Dynamics Review 17(4): 347-356.

Hudson, M., Smart, A. and Bourne, M. 2001. Theory and practice in SME performance measurement systems. International Journal of Operations and Production Management 21: 1096-1115.

Kaplan, R.S. and Norton, D.P. 1992. The balanced scorecard: measures that drive performance. Harvard Business Review 70: 71-79.

Kaplan, R.S. and Norton, D.P. 2001. Leading change with the balanced scorecard. Financial Executive 17: 64-66.

Lane, D.C. 1992. Modelling as learning: A consultancy methodology for enhancing learning in management teams. European Journal of Operational Research 59(1): 64-84.

Lohman, C., Fortuin, L. Wouters, M. 2002. Performance management as a sporty exercise. Beta Research Working Paper W69, Eindhoven University of Technology, Eindhoven.

Malmi, T. 2001. Balanced scorecards in Finnish companies: a research note. Management Accounting Research 12: 207-220.

Martinsons, M., Davison, R. and Tse, D. 1999. The balanced scorecard: a foundation for the strategic management of information systems. Decision Support Systems 25: 71-88.

Miller, G. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review 63: 81-96.

Mooraj, S., Oyon, D. and Hostettler D. 1999. The balanced scorecard: a necessary good or an unnecessary evil. European Management Journal 17: 481-491.

Morecroft, J.D.W. 1983. System dynamics. Portraying bounded rationality. OMEGA 11(2): 131-142.

Morecroft, J.D.W. 1984. Strategy support models. Strategic Management Journal 5(3): 131-229.

Morecroft, J.D.W. 1985. Rationality in the analysis of behavioral simulation models. Management Science 31(7): 900-916.

Neely, A., Gregory, M. and Platts, K. 1995. Performance measurement system design. International Journal of Operations and Production Management 15: 80-116.

Neely, A. 1998. Measuring Business Performance. Profile Books Ltd: London.

Newing, R. 1994. Benefits of a balanced scorecard. Accountancy November: 52-53.

Nørreklit, H. 2000. The balance on the balanced scorecard – a critical analysis of some of its assumptions. Management Accounting Research 11: 65-88.

Paich, M. and Sterman, J.S. 1993. Boom, bust and failures to learn in experimental markets. Management Science 39(12): 1439-1458.

Platts, K.W., Tan, K.H. 2002. Designing linked performance measures – A connectance based approach. Pre-prints 12th Intern. Seminar on Production Economics, Vol. 2: 367-373.

Pidd, M. 1996. Tools for thinking. Modelling in management science. Wiley, Chichester.

Pugh, D.S. and Hickson, D.J. 1989. Writers on Organisations. 4th ed. Penguin Books: London.

Richardson, G.P. and Andersen, D.F. 1995. Teamwork in group model building. System Dynamics Review 11(2): 113-137.

Richmond, B. 1994. Systems thinking/system dynamics: let's just get on with it. System Dynamics Review 10(2-3): 135-57.

Senge, P. 1990. The fifth discipline. The art and practice of the learning organisation. Doubleday Currency, New York.

Siegele, L. 2002. How about now? A survey of the real-time economy. The Economist, February 2 2002: 3-5.

Simon, H. 1957. Administrative Behavior: a Study of Decision-Making Processes in Administrative Organisations. 2nd ed. Macmillan: New York.

Sterman, J.D. 1989. Modelling managerial behavior. Misperceptions of feedback in a dynamic decision making experiment. Management Science 35(3): 321-339.

Sterman, J.S. 2000. Business dynamics. Systems thinking and modelling for a complex world. McGraw-Hill, New York.

Vennix, J.A.M. 1996. Group model building. Facilitating team learning using system dynamics. Wiley, Chichester.

Warren, K. and Langley, P. 1999. The effective communication of system dynamics to improve insight and learning in management education. Journal of the Operational Research Society 50: 396-404.

Winch, G.W. 1993. Consensus building in the planning process: benefits from a "hard" modelling approach". System dynamics review 9(3): 287-300.

Wisniewski, M., Dickson, A. 2001. Measuring performance in Dumfries and Galloway Constabulary with the Balanced Scorecard. Journal of the Operational Research Society 52: 1057-1066.

Wolstenholme, E.F. and Coyle, R.G. (1983) The development of system dynamics as a methodology for system description and qualitative analysis. Journal of the Operational Research Society 7: 569-581.

Wolstenholme, E.F. 1999. Qualitative vs quantitative modelling: the evolving balance. Journal of the Operational Research Society 50: 422-428.

Yerkes, R. and Dodson, J. 1908. The relation of strength of stimulus to rapidity of habit formation. Journal of Comparative Neurology and Psychology 18: 459-482.

Yin, R.K. 1989. Case Study Research: Design and Methods. Sage, London.

Back to the Top