# A system dynamics-based simulation experiment for testing mental model and performance effects of using the balanced scorecard

Carlos Capelo<sup>abc</sup>\* and João Ferreira Dias<sup>d</sup>

Carlos Capelo holds a PhD from ISCTE in Strategic Management. He is Professor of Management at Universidade Lusófona and Instituto Superior de Gestão, Lisbon, Portugal. He has also been developing a management career in Galp Energia, a Portuguese oil & gas company. His main research areas are related to strategic learning and business modelling and simulation.

João Ferreira Dias holds a PhD from ISCTE in Management. He is Professor of Management at ISCTE —Instituto Superior de Ciências do Trabalho e Empresa, Lisbon, Portugal. He is also a Senior Researcher at ADETTI. His main research areas are related to strategic management and system dynamics.

#### Abstract

This study develops a theoretical model that explains the effectiveness of the balanced scorecard approach by means of a system dynamics and feedback learning perspective. Presumably, the balanced scorecard leads to a better understanding of context, allowing managers to externalize and improve their mental models. We present a set of hypotheses about the influence of the balanced scorecard approach on mental models and performance. A test based on a simulation experiment that uses a system dynamics model is performed. The experiment included three types of parameters: financial indicators; balanced scorecard indicators; and balanced scorecard indicators with the aid of a strategy map review. Two out of the three hypotheses were confirmed. It was concluded that a strategy map review positively influences mental model similarity, and mental model similarity positively influences performance. Copyright © 2009 John Wiley & Sons, Ltd.

Syst. Dyn. Rev. 25, 1-34, (2009)

# Improving learning and performance with the balanced scorecard

# Mental models, double-loop learning, cognitive limitations, quality of feedback, and performance

A mental model is a conceptual representation of the structure of an external system used by people to describe, explain and predict a system's behavior (Craik, 1943; Johnson-Laird, 1983). Mental models have been commonly used in system dynamics and systems thinking literature (Forrester, 1961; Senge, 1990; Doyle and Ford, 1998, 1999; Sterman, 2000). Managers build their mental models as they interact with the business system that they manage. Experimental research has suggested that decision makers perform better if the structure of their mental model is as similar as possible to the structure of the external system it represents (Rowe and Cooke, 1995; Stout *et al.*, 1996; Wyman and Randel, 1998; Davis and Yi, 2004; Mathieu *et al.*, 2000, 2005). In particular, experiments conducted within the system dynamics community (Ritchie-Dunham, 2001, 2002; Ritchie-Dunham *et al.*, 2007; Gary and Wood, 2007) using interactive

<sup>a</sup> Faculdade de Economia e Gestão, Universidade Lusófona, Campo Grande 376, Lisbon, 1749-024, Portugal.

\* Correspondence to: Carlos Capelo. E-mail: carlos.capelo@galpenergia.com Received June 2006; Accepted September 2008

System Dynamics Review Vol. 25, No. 1, (2009): 1–34 Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/sdr.413 Copyright © 2009 John Wiley & Sons, Ltd.

<sup>&</sup>lt;sup>b</sup> Dep. Gestão, Instituto Superior de Gestão, Lisbon, Portugal.

<sup>&</sup>lt;sup>c</sup> GPL, Galp Energia, Lisbon, Portugal.

<sup>&</sup>lt;sup>d</sup> ADETTI, ISCTE, Lisbon, Portugal.

computer-based simulations have shown that the accuracy of a person's mental model is a good predictor of their performance. Thus, it is reasonable to assume that, by enhancing their mental models, managers should be able to improve their capacity to deal with dynamically complex relationships in both internal and external environments. This in turn would improve their ability to manage the business system in order to achieve long-term success.

Managers make decisions and learn in the context of feedback loops (Forrester, 1961). In single-loop learning, managers compare information about the state of a real system to pre-established goals, perceive deviations between desired and actual states, and make the decisions they believe will move the system towards the desired state. Single-loop learning does not change the managers' mental models. In double-loop learning, information about the business system is not only used to make decisions within the context of existing frames, but also feeds back to modify the managers' mental models (Argyris, 1976). As their mental models change, managers define new strategies and policies.

Cognitive limitations and quality of feedback information are key factors that impact the effectiveness of double-loop learning because they have the potential to limit managers' perception and understanding of the actual business system (Richardson et al., 1994; Sterman, 2000). Owing to cognitive limitations, the mental models that managers use for decision making are necessarily imperfect, with the result that a flawless assessment of the dynamic behavior of the business system is nearly impossible (Sterman, 2000). Strategic learning processes are also strongly influenced by the quality of information fed back about the state of the business system. Managers use that information to interact with the business system. In the event of imperfect feedback information, managers have an incorrect perception of the impact of their decisions, and so they are unable to build their mental models accurately (Sterman, 2000). Nevertheless, the double-loop process, although significantly impaired by possible cognitive limitations and imperfect feedback information, offers enormous potential for effective learning if means can be developed to overcome these limitations. Thus, appropriate performance measurement systems must be designed and implemented in order to overcome or minimize these barriers to strategic learning.

#### The balanced scorecard approach

Kaplan and Norton (1992, 1996a, 2001a) introduced the balanced scorecard (BSC) with the aim of overcoming strategic management limitations of the traditional performance measurement systems, as they were too narrowly focused on financial measures. The BSC approach features a mix of lead (performance drivers) and lag (outcome measures) indicators; financial and non-financial measures; and hard and soft, more subjective measures. These are categorized as follows: (i) financial; (ii) customer; (iii) process; and (iv) learning and growth. The BSC tool helps managers monitor actual financial and market performance, evaluate

the results of short-term processing actions, monitor the intangible development of competencies that will drive future financial performance and assess the progress of implementing corporate strategy. According to the authors, the BSC approach provides top managers with a picture of a possible future (a vision), a path for getting there (a strategy) and a mapping to medium- and short-term quantifiable objectives and actions. Strategy implies the movement of an organization from its present state to a desirable but uncertain future state. Since the organization has never been in such a future state, the pathway to it is built on a series of linked hypotheses (Kaplan and Norton, 2000). To support managers in developing a cause-and-effect perspective and to better understand the business system in which they participate, Kaplan and Norton (2000, 2001a) developed the strategy map concept as a complementary tool to the BSC approach. The strategy map links the performance indicators in a causal chain (causal diagram) that helps managers to translate, test and communicate their understanding of the business system. This is a valuable step towards implementing and revising company strategy. Kaplan and Norton (2000) define a strategy map as a visual representation of a company's critical objectives and the crucial relationships among them that drive organizational performance. A more recent study (Kaplan and Norton, 2004) states that "We now realize that the strategy map, a visual representation of the cause-and-effect relationships among the components of an organization's strategy, is as big an insight to executives as the balanced scorecard itself."

It is clear that the BSC approach is consistent with the systemic and dynamic view of business management and performance measurement. The BSC framework recognizes the interconnectedness within the business and the importance of understanding cause-and-effect relationships and their dynamics as a consistent basis upon which to infer future performance and define objectives and action plans (Warren, 2002; Ritchie-Dunham and Puente, 2008). Strategy maps combined with balanced scorecards provide an integrated and holistic approach to business management and performance measurement. The strategy map describes the managerial perception of the structure of the business system, while the performance measurement information provided by the BSC captures the essential nature of the system's behavior. The strategy map can be regarded as a systems thinking tool for modeling strategy. From this perspective, Kaplan and Norton even suggested that "the BSC can be captured in a system dynamics model that provides a comprehensive, quantified model of a business's value creation process" (Kaplan and Norton, 1996b).

# Improving learning and performance with the balanced scorecard

In the BSC framework, strategies are seen as hypotheses. Managers should be able to test, validate and review these hypotheses. The BSC provides feedback information to managers, allowing a better understanding of the business system and strategic redesign. Kaplan and Norton (2001a) argue that the BSC



approach supports double-loop learning that facilitates managerial strategic learning, leading to better performance. Kaplan and Norton (2001b) describe the process of strategic learning and adaptation, as presented in Figure 1. The pathway includes three components: organizations use BSC to link strategy to the budgeting process; management meetings to review strategy are introduced; and finally a process for learning and strategy adaptation evolves. In a continual process, managers use the balanced scorecard and the strategy map to re-evaluate the assumptions used in the previous strategy. They review the assumed cause-and-effect relationships and identify new ones. Then they improve their understanding of the business system and they determine a new strategy (Kaplan and Norton, 2001a). In other words, BSC triggers a process by which managers can make explicit improvements to their mental models of the business system. They adapt the company strategy and define the new short- and medium-term objectives by simulating their mental models to infer the future behavior of the business system.

#### Balanced scorecard and organizational performance

Since BSC was firstly introduced (Kaplan and Norton, 1992), an enormous number of books and articles that describe and recommend BSC have been published. In that literature, one can find comments such as: the BSC is "among the most significant developments in management accounting" (Atkinson *et al.*, 1997); the BSC "is a necessary good for today's organizations" (Mooraj *et al.*, 1999); the BSC "provides a new foundation for strategic control" (Olve *et al.*, 2000); "Scorecards and strategy maps will be used as an organization's common language for discussing the rationale behind actions" (Olve *et al.*, 2003). According to Kaplan and Norton (2001a), many organizations around the world are using the BSC approach to define, implement and manage strategy. In fact, recent surveys showed that BSC was the most popular performance measurement system, adopted by over 40 percent of organizations worldwide (57 percent in the UK, 46 percent in the USA, 28 percent in Germany and Austria) (Rigby, 2001; Speckbacher *et al.*, 2003).

Companies around the world continue making significant investments in the development and implementation of BSC systems. Considering that these investments are based on the hypothesis that the use of BSC has a positive impact on the performance of the organization, it is important to investigate whether these systems lead to an improvement in strategic learning and decision effectiveness. However, empirical research into the performance implications of the BSC approach is still scarce. And the positive contributions of the BSC have not been unambiguously confirmed by previous studies. There exist studies in the field of management accounting research that identify problems and limitations associated with the BSC approach. The inadequate definition and utilization of the performance indicators have been highlighted as a main drawback of the BSC system (Lingle and Schiemann, 1996; Stivers et al., 1998; Ittner and Larcker, 1998; Lipe and Salterio, 2000; Malmi, 2001; Speckbacher et al., 2003). Those studies show in general terms that the measures and perspectives in use are fairly independent, and do not always mirror the recommended cause-and-effect logic of the BSC approach. Ittner and Larcker (2003) reported that only 23 percent of the 157 organizations surveyed consistently built and tested causal models to support the definition of their performance indicators, but that these organizations on average did achieve a superior level of performance. Ittner et al. (2003) and Hendricks et al. (2004) found no meaningful performance improvements as a result of BSC usage. Braam and Nijssen (2004) found that BSC use that complements company strategy leads to higher performance, while the performance effect of measurement-focused BSC use was significantly negative. Lipe and Salterio (2000) carried out an experiment to examine the judgmental effects of BSC in a business unit and found that evaluators frequently place greater or exclusive emphasis on common measures and tend to under-weight the influence of unique measures. Common measures tend to be lagging and financial, while unique measures are more often leading and non-financial (Kaplan and Norton, 1996a). Consequently, this research suggests that, due to cognitive limitations, managers may not appreciate the significance of non-financial and leading measures, thereby defying the cause-and-effect logic built into the BSC approach and potentially reducing the benefits of using the BSC. Akkermans and van Oorschot (2005) describe a case study where a two-stage modeling process (qualitative causal loop diagramming followed by quantitative simulation) was used to overcome certain problems associated with BSC implementation. This research suggests that qualitative causal loop diagramming helps managers in identifying key variables and their causal interrelations, and the use of system dynamics simulation modeling is essential in developing a better comprehension of business dynamics, such as time delays and accumulations in the key business processes.

Some simulation-based experiments have been carried out within the system dynamics community with the aim of testing the impact of BSC on performance. In one simulation-based study (Ritchie-Dunham, 2001, 2002; Ritchie-Dunham *et al.*, 2007), where subjects were asked to manage a firm by interacting with a system dynamics-based micro-world, it was found that similarity between the subjects' mental models and the structure of the simulation model positively impacted the performance improvement associated with BSC usage. These results also showed that BSC use led to mental models that more closely mirrored reality, with a commensurate positive impact on performance. The simulation-based experiments reported by Strohhecker (2007) also found evidence of a positive impact of BSC on performance.

#### **Research model**

This research focuses on how the BSC approach facilitates strategy review and implementation, leading to enhanced double-loop learning effectiveness, and how this type of learning influences performance. The expected relations and hypotheses, based on the following variables, are pictured in Figure 2.

- Level of Balanced Scorecard (LBSC). This variable represents the extent of BSC use as a performance measurement system. In other words, this indicates to what extent managers focus on a BSC structure of performance indicators that combine critical financial and non-financial measures, instead of only relying on financial indicators when taking decisions.
- *Level of Strategy Map* (LSM). This variable represents the intensity of strategy map use as a tool associated with the BSC approach to support the process of strategy review and implementation.





- *Mental Model Similarity* (MMS). This variable represents the participants' understanding of the structure of the business system, and measures the similarity between the structure of managers' mental models and the structure of the business system as represented in the simulator.
- *Performance*. The performance of this management task is measured in terms of the financial value created by the firm.

The performance measurement system associated with the BSC approach is viewed as a comprehensive structure of performance indicators that combine both financial and non-financial measures. The goal is to capture and understand the core elements of the business system's behavior. Consistent with the assumptions of the BSC approach, in order to make more effective decisions managers should focus on a mix of financial and non-financial indicators, instead of relying only on financial indicators. From a feedback system perspective, this means that managers who only use financial indicators (low LBSC) tend to be unable to learn and act effectively, as they are not aware and do not access the behavior of other critical components within the business system. A higher LBSC represents more comprehensive use of the performance measurement system associated with the BSC approach. This provides managers with more effective feedback information about business behavior that drives future performance. By using a higher LBSC, managers access more relevant information to interact with the business system, and they acquire a sharper understanding of the impact of their decisions. This relation improves their ability to build more accurate mental models, leading to higher MMS. In other words, the present work assumes that if managers use the BSC performance measurement system in the process of strategic review and implementation, they generate more effective double-loop learning and, consequently, more appropriate mental models.

*Hypothesis 1*: The use of the BSC performance measurement system is positively correlated with MMS.

It is assumed that by using the BSC strategy map tool managers will benefit from more effective double-loop learning as they review the critical causeand-effect relations through a process that externalizes and improves their mental models of the business system.

*Hypothesis 2*: The use of the BSC strategy map tool is positively correlated with MMS.

This research assumes that managers perform better if the structure of their mental model is as similar as possible to the structure of the external system it represents.

*Hypothesis 3*: MMS positively influences Performance (financial value creation).

#### Method

System dynamics based simulation models have increasingly attracted the attention of researchers within this field. System dynamics appears to be an effective and relevant tool to create underlying formal simulation models for research purposes (Größler, 2001; Repenning, 2003). In several areas of management research, computer simulators based on system dynamics models are used as a means to explore the subjects' understanding and behavior in complex situations. In particular, within the system dynamics community they are well accepted and frequently used as instruments for investigating human cognition and decision making in complex business situations (Sengupta and Abdel-Hamid, 1993; Paich and Sterman, 1993; Sterman, 1989; Howie *et al.*, 2000).

We considered a system dynamics model to be an appropriate research tool for our problem. The hypotheses defined in this research and presented in the previous section were tested with a simulation-based experiment in which subjects interacted with a system dynamics micro world that provided information through a BSC interface. This section presents an overview of the chosen simulator, describes the subjects and the experiment conditions, and overviews the research model variables.

#### Simulator overview

The business simulator was built by incorporating the same system dynamics model that had been used in previous research (Ritchie-Dunham, 2001, 2002; Ritchie-Dunham *et al.*, 2007). This study uses the same business case, model structure, game interfaces and initial conditions that were defined by Ritchie-Dunham (2002). The text and simulator interfaces were translated into Portuguese. Apart from the translation, the model used in this experiment is identical to the original model.

The participants run a realistic simulator of a wireless telecommunications firm by making critical decisions every 6 months for a simulation period of 7 years. They analyze the business status and make the following decisions: investing financial resources in (1) infrastructure (base stations) and (2) information technology; and decisions on human resources in terms of (3) training, (4) hiring, and (5) firing employees. These resources (Base Stations; Information Technology and Skilled Employees) influence different dimensions of Customer Satisfaction, which influences the number of customers in the Customer Base and has an impact on the organization's financials. The participants analyze the feedback information from financial and non-financial indicators in order to determine further investment and human resource decisions before the cycle repeats.

Figure 3 presents a simplified representation of the stock and flow model of the simulator, showing how the critical resources and other variables are interrelated



Fig. 3. Simplified representation of the simulator model

and its association with the balanced scorecard perspectives (decision variables are shown in bold).

- *Growth and learning perspective*. The level of information technology (IT), which influences the IT Facilitation Index, is accumulated when the firm invests in new IT and is withdrawn at a specific time thereafter. Employees have either relevant skills or obsolete skills. Skilled Employees are hired and may leave the firm by attrition or downsizing. They are converted to Obsolete Employees if their skills are made obsolete by newer information technologies. Obsolete Employees leave by attrition or downsizing, or may be trained to become Skilled Employees. The rate of employees being trained is influenced by training effectiveness, which in turn is influenced by the annual training budget. The HR (Human Resource) Service Index is influenced by the ratio of actual customers to skilled employees.
- Business process perspective. The firm decides to invest in infrastructure (Base Station Investment) and places orders for base stations with the suppliers, which accumulates the stock of "Base Stations in Process" (BSIP); in the simulator this refers to Base Stations that are in the process of being acquired. The suppliers' daily building capacity available to the firm is a function of the suppliers' annual building capacity and the firm's market share. This supplier building capacity is adjusted due to changes in demand and changes in the firm's output. The response time of the supplier's growth in building capacity causes stock accumulation of BSIP and delays the construction of infrastructure. Base stations eventually get converted from BSIP into Base Stations (in service). After a certain average usable lifetime, the Base Stations are closed. The number of Base Stations and the number of Customers determine the Capacity Utilization which influences Network Quality, which in turn influences the customers' Perceived Call Quality. The perceived call quality is also influenced by Network Coverage. Perceived Customer Service is a function of the IT Facilitation Index and the HR Service Index.
- *Customer perspective*. Perceived Call Quality determines Customer Satisfaction, which is also influenced by Perceived Customer Service. Customer Satisfaction influences the firm's Retention Rate, which influences Customer Flows.
- *Financial perspective*. Some financial variables (not shown in Figure 3) and related to revenues, costs, capital and value creation are calculated from the status and flow of resources in the firm.

A more detailed description of the business case and simulator model can be found in Ritchie-Dunham (2002) and Ritchie-Dunham *et al.* (2007).

The initial conditions and model structure were the same for all participants. The participant objective was to develop those critical and interrelated resources (Base Stations, IT, Skilled Employees, and Customers) at appropriate rates and levels in order to gain and retain customers, operate efficiently, and



Fig. 4. Simulator interface for financial scorecard (panel control with initial values). (Adapted from Ritchie-Dunham, 2002)

maximize value creation. To succeed in this simulation task, participants had to identify and understand the cause-and-effect relationships among critical variables. As all resources had to be consistently developed, participants also needed to recognize and address both delay and stock accumulation effects, especially those related to the process of building infrastructure (due to the response time associated with the supplier's need to increase building capacity).

The simulator provides two alternative interfaces. One represents a financial scorecard interface (Figure 4), which features EBIT and other measures that are directly related to its calculation (Ritchie-Dunham, 2002). The other interface represents a BSC (Figure 5), which includes a set of leading and lagging, financial and non-financial indicators that are graphically separated into four sections related to the four perspectives associated with the BSC approach (Ritchie-Dunham, 2002). Each simulator interface includes three screens: the first screen allows participants to enter their decisions and provides data for that time period; the second screen presents the historical behavior over time for each of the variables in the first screen; the third screen provides a description of each of the variables in use.



Fig. 5. Simulator interface for balanced scorecard (panel control with initial values). (Adapted from Ritchie-Dunham, 2002)

# Subjects, apparatus, procedure, and facilitation

This research was conducted at ISCTE (a business graduate school in Lisbon) and at Galp Energia, a Portuguese oil company and one of the largest firms in the country. At ISCTE the group consisted of 14 undergraduate students in their last year of a 4-year business degree. Their ages ranged from 22 to 25 and they had no work experience. At Galp Energia the task was performed by a group of 59 managers. Their ages ranged from 25 to 54 and on average they each had 13 years of work experience. The simulation task was performed individually (participants could not interact), anonymously and without financial rewards or incentives. As articulated by Camerer and Hogarth (1999), there is no evidence that financial incentives improve performance in these types of experimental tasks. Moreover, the subjects were highly motivated to perform the task as they were interested in exploring the BSC and had volunteered to participate in the simulation experiment. All the participants were familiar with basic BSC concepts and with the financial measures used to calculate and define task performance. The participants had no experience

with the simulator and they also had no prior specific knowledge about wireless telecommunications businesses.

At ISCTE, the experiment was carried out in a computer laboratory with one participant per computer. At Galp Energia, each participant performed the simulation on his/her own office computer. All participants were given a full experimental guide including: description and objective of the simulation task; case text; instructions for accessing and starting the simulator on the computer network; instructions for running the simulator; questionnaire about strategy and objectives; sheets for strategy map review (only for participants using strategy map); and a questionnaire about the relatedness of certain simulator variables.

The decisions made on the simulation and its results were automatically stored in a protected spreadsheet on the participant's computer. Each participant took about 120 minutes to perform the task.

In the simulation experiment, the participants were involved in the dynamical decision-making processes presented in Figure 6. They analyzed business status using the simulator interface, used this information to review the strategy and objectives and decision making, and then repeated the process. There were three different stages. In stage A the participants ran the firm by using the financial scorecard interface (Figure 4); in stage B the firm was operated using the BSC interface (Figure 5); and in stage C the firm was run by using the BSC and the strategy map.



Fig. 6. Stages and dynamic decisionmaking process procedures



As shown in Figures 6 and 7, stages A and B had the same procedure and involved just one session. The procedure for stage C was different, as participants reviewed the strategy map and participated in two sessions. The experimental procedures had the following common initial steps. Firstly, the participants were randomly assigned to one of three stages (A, B or C). Next, they answered some questions (age, simulation experience, and management experience), and they read the introduction with the overall description and the objectives of the simulation and the business case study. Finally, the participants were instructed to raise any questions they had as they proceeded with the case.

For stages A and B, in the same session, the participants read the instructions for accessing, starting and running the simulator, and they were given oral instructions with examples to show simulator operation. A first simulation was conducted to familiarize participants with the game interfaces and commands. The actual experiment involved a second simulation that included a questionnaire about strategy and objectives every 2 years of simulation time. After the simulation, the participants answered a questionnaire about their understanding of the linkages between certain critical concepts.

The first session of stage C called for participants to read the business case text before performing the following steps. Firstly, the participants filled out the same questionnaire that was used on the final step of stages A and B about their ultimate understanding of the linkages between certain critical variables-this questionnaire captured their initial level of comprehension of the business system. The answers to that questionnaire were automatically translated into a network diagram using the Pathfinder procedure (Schvaneveldt, 1990; Rowe and Cooke, 1995). The second session commenced with the analysis of an initial strategy map, drawn from the previous network diagram, that showed the linkages among critical concepts (example in Figure 9). Participants received instruction on how to read and interpret the initial strategy map, and how to review it by cutting or inserting links between the indicators and defining the arrows that indicated the cause-and-effect relationships. Since the initial strategy map only showed linkages among concepts, they were encouraged to draw arrows to define the cause-and-effect relationships among those variables (example in Figure 10). Next, they read the instructions for accessing, starting and running the simulator. They ran a first simulation to familiarize themselves with the game interfaces and commands, and at this point they were instructed to ask for help at any time. Finally the participants performed the simulation that represented the actual experiment; during stages A and B, they answered a questionnaire about strategy and objectives every 2 years of simulation time. However, this experimental group was also asked to review the causal diagram (strategy map) which had been provided in the experiment guide. They cut or inserted links so that the causal diagram accurately expressed their latest understanding of the simulated business system. The participants were also encouraged to use the strategy map to reflect on strategy, objectives, and decisions. Finally, they drafted a final strategy map—this map represented their ultimate understanding of the business system.

# Research model variables

This section summarizes the use of the variables LBSC, LSM, MMS, and Performance, which were defined in the research model. In order to perform some exploratory analysis, the following variables were also defined and measured: Time (total time participants spent on the task); Age (participant's age); Simulation Experience (previous experience with management simulators); and Work Experience (previous work experience in management).

- *LBSC*. The variable LBSC features two degrees. In the low degree (low LBSC), the subjects ran the firm only by using the financial indicators included in the financial scorecard interface (stage A); in the high degree (high LBSC), the subjects ran the firm using the BSC indicators included in the BSC interface (stages B and C).
- *LSM*. The variable LSM also features two degrees. In the low degree (low LSM), the participants ignore the strategy map while running the simulator (stages A and B); in the high degree (high LSM), the strategy map is used to define and review the strategy and objectives (stage C). The strategy map (examples

in Figures 10 and 15) that is used throughout this experiment consists of a causal diagram containing simulator concepts spatially organized into four sections related to the four perspectives of the BSC. The strategy map utilization is described in the previous section (under stage C procedures).

• *MMS*. The variable MMS represents the participants' understanding of the structure of the business system. As the structure of the simulated business system is known by the researchers in advance, it can be compared with the participants' mental model in order to evaluate how that elicited mental model matched the simulated reality. This variable measures the similarity between the structure of the subjects' mental models and the representative structure of the simulator (Figure 8) (Rowe and Cooke, 1995; Ritchie-Dunham, 2002).



Fig. 8. The representative network of the simulated business system. The decision-maker evaluates the information from critical indicators such as Perceived Call Quality, Perceived Customer Service, Customer Satisfaction, and Economic Value Added, and makes further decisions by acting on IT investment, HR Hiring Rate, HR Training Investment and Base Stations Investment. These decisions result in changes to Perceived Customer Service and Perceived Call Quality, which influence Customer Satisfaction and Revenues. All investment decisions influence Capital Cost. Economic Value Added is calculated from Revenues, Capital Cost and Total Operating Cost



Fig. 9. Example of initial strategy map (participant Z from group C)

In the lower level of the strategy map (defined in connection with stages A and B in the previous section), participants were asked to fill out a questionnaire detailing their final understanding of the simulated business system. This questionnaire used a nine-point scale to evaluate the relatedness of 14 concepts in the simulation model (Rowe and Cooke, 1995; Ritchie-Dunham, 2002). The 14 concepts are relevant to understanding the simulated business system. The 91 pairings needed to relate the 14 concepts  $(n^2/2 - n/2 = 14^2/2 - 14/2 = 91)$  were presented in random order. The structure of the participant's mental model could be elicited by this pair-wise relatedness ratings technique. These elicited pairings were transformed into a network diagram using a network scaling procedure pathfinder (Schvaneveldt, 1990). In the higher LSM (defined in the previous section under the discussion of stage C) the participants produced a final strategy map linking the same 14 concepts as were found in the simulation model. This final strategy map represented the elicited structure of the subjects' mental model.



Fig. 10. Example of a final strategy map (causal diagram), drafted by the participant who was given the initial strategy map in Figure 9 (participant Z from group C)

MMS was measured in terms of the similarity between two networks: the participants' mental model and the representative network of the simulator. This network similarity ranged from 0 (low similarity) to 1 (high similarity) and was determined by the number of common links divided by the total number of links in both networks (Schvaneveldt, 1990; Rowe and Cooke, 1995; Ritchie-Dunham, 2002).

• *Performance*. Task performance was measured by total financial value creation. This value was estimated by summing the discounted economic profit or economic value added (defined as Net Operating Profit Less Amortizations and Taxes – Weighted Average Cost of Capital × Total Capital Employed) of the firm over the seven simulated years (Copeland *et al.*, 2000). The performance measurement only took into account the economic value added; the firm's future value was not taken into consideration. This condition was defined so that the task objectives were made clear, allowing the participants

to focus on the simulation period (7 years). Moreover, previous experience with this simulator had shown that the final decisions made by each participant were strong predictors of future value.

# **Results and discussion**

The 73 participants were grouped as follows:

- *stage A*: using financial scorecard interface without strategy map (low LBSC, low LSM)—24 participants;
- *stage B*: using BSC interface without strategy map (high LBSC, low LSM)—24 participants;
- *stage C*: using BSC interface and strategy map (high LBSC, high LSM)—25 participants.

Table 1 presents minimum, maximum and mean values, and standard deviations for the dependent variables of each stage group. Table 2 shows the results of statistical testing to identify differences in means between the stage groups.

The participants in group C—BSC interface and strategy map review—showed on average the best MMS (mean = 0.443) and the best Performance (mean = 628). As shown in Table 2, the mean values of MMS and Performance for group C were significantly different from the equivalent values for groups A (mean difference = 0.189, p < 0.001) and B (mean difference = 0.144, p < 0.001). On average, the participants of group B—BSC interface—showed a better MMS (mean = 0.295) than those in group A—financial scorecard interface (mean = 0.250). Table 2 shows that the difference is significant at p < 0.05 (mean difference = 0.045, p = 0.043). Participants of group A (mean = 329) and participants of group B (mean = 310) showed a similar mean value for Performance (mean difference = 18, p = 0.925).

The lowest values for the variables MMS and Performance were found in participants from group B—BSC interface (Table 1). This result may be explained by the overload caused by an excessive amount of information—much larger than the information volume that was processed by participants of group A

		Mental Mo	Mental Model Similarity			Performance		
Stage	Description	Min/Max	Mean	SD	Min/Max	Mean	SD	
А	Low LBSC, Low LSM	0.122/0.406	0.250	0.080	-715/854	329	450	
В	High LBSC, Low LSM	0.093/0.429	0.295	0.077	-1148/1189	310	687	
С	High LBSC, High LSM	0.205/0.708	0.443	0.126	-432/1089	628	409	
Cbs	Before simulation	0.128/0.442	0.253	0.089				

Table 1. Means and standard deviations for Mental Model Similarity and Performance for each stage group

> Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/sdr

Table 2. Tests of significance for differences in means between the stage groups

Pair	Mental Model Similarity			Performance		
	Mean Difference	SD	Significance p	Mean Difference	SD	Significance p
A–B	-0.045**	0.102	0.043	18	939	0.925
B–C	-0.144***	0.153	0.000	-313**	632	0.023
A–C	-0.189***	0.139	0.000	-295**	592	0.023
A–Cbs	0.002	0.118	0.950			
B–Cbs	0.043*	0.121	0.093			
C–Cbs	0.190***	0.135	0.000			

p < 0.1; p < 0.05; p < 0.001.

with the financial scorecard interface. This may have led participants to misinterpret the indicators of structure and behavior.

In Table 1, the mean value for MMS for stage C before simulation (Cbs) represents the participants' understanding of the simulated business system after reading the case text. Table 2 shows that the mean difference for MMS between stage C and Cbs is very significant (mean difference = 0.190, p < 0.001). It is clear that the differences in means for MMS from participants in groups A and B compared with participants in group C—before simulation— are not very significant. This suggests that, on average, participants from group A (using financial scorecard) and B (using BSC without strategy map) did not learn much about the simulated business system.

The results suggest that strategy map review gave participants in group C a significant cognitive aid that accelerated their learning about the simulated business system. This simulation experiment did not involve a formal briefing session. However, after being informed about the global results, some participants of group C argued that, by accessing and reviewing the initial strategy map just after finishing the practice simulation, they tested their initial assumptions more effectively. Therefore, they may have benefited right from the start of the simulation that constituted the actual experiment with a better understanding of the simulator.

As an illustrative example of strategy map review, the initial and final diagrams drawn by a participant from group C (designated as "Z") are presented in Figures 9 and 10. Before starting the simulation, Z was given the initial strategy map shown in Figure 9. He interpreted the diagram and defined the cause– effect relationships. Then, he continuously reviewed the diagram during the simulation by inserting and cutting links (arrows). The final strategy map drafted by Z is presented in Figure 10. Some common behaviors revealed in the simulation experiment can be observed in this example (Figures 9 and 10). Links that were originally in the network were eliminated, yielding a simpler and clearer diagram. The links among financial indicators were correctly reviewed (revenues, operating cost, capital cost, and economic value added).

		Dependen	t variables	formance			
	Mental Mode	el Similarity	Performance				
Independent variables	Standardized beta	Significance p	Standardized beta	Significance p			
Time	-0.011	0.918	0.038	0.732			
Age	0.011	0.926	-0.271**	0.039			
Work Experience	0.179	0.152	0.274**	0.043			
Simulation	-0.150	0.168	0.159	0.176			
Experience							
LBSC	0.227*	0.050	-0.032	0.795			
LSM	0.542***	0.000	-0.092	0.502			
MMS			0.595***	0.000			
Adjusted R <sup>2</sup>	0.397		0.319				

Table 3. Regression results for all independent variables

p < 0.1; p < 0.05; p < 0.005; p < 0.001.

The representative network of the simulated business system that was used to calculate MMS values considers some feedback loops which represent the interaction of the decision-maker (participant) acting within the business system. As can be seen in Figure 10, Z did not draw those feedback loops on the strategy map. He considered a linear causal net starting with the four decision variables (IT Investment, Base Stations Investment, HR Hiring Rate and HR Training Investment) and ending at the Economic Value Added variable.

Table 3 shows the results of multivariate regression analyses of MMS and Performance on the independent variables. The regressions were run on standardized values for all the variables so as to directly compare the relative effects of each independent variable on the dependent variable. As presented in Table 3, regression analysis for MMS on the independent variables shows a highly significant effect for LSM ( $\beta = 0.542$ , p < 0.001), a significant effect for LBSC ( $\beta = 0.227$ , p = 0.050) and no significant effects for other variables. Regression analysis of Performance shows a highly significant positive effect for MMS ( $\beta = 0.595$ , p < 0.001), a significant positive effect for MMS ( $\beta = 0.274$ , p = 0.043), a significant negative effect for Age ( $\beta = -0.271$ , p = 0.039), and no significant effects for other variables.

The regression model was refined by performing a stepwise regression in order to exclude the variables that did not seem to significantly explain the dependent variables and to preserve the most significant explanatory variables (Figure 11). As presented in Figure 11, regression analysis of MMS on the most significant independent variables shows a very strong effect for LSM ( $\beta = 0.529, p < 0.001$ ) and low significance for LBSC ( $\beta = 0.191, p < 0.1$ ). Regression analysis of Performance shows a very significant effect for MMS ( $\beta = 0.528, p < 0.001$ ), a significant positive effect for Work Experience ( $\beta = 0.343, p < 0.01$ ) and a significant negative effect for Age ( $\beta = -0.237, p = 0.1$ ).





Unexpectedly, the results indicate that the variables Time (total time participants spent on the task) and Simulation Experience (previous experience in business game simulators) did not influence MMS or Performance. Interestingly, the findings show a positive effect of Work Experience on Performance, but this variable does not seem to influence MMS. In fact, the Work Experience variable differentiates two specific groups of participants: the managers from Galp Energia (with an average of 13 years of work experience) and the students from ISCTE (without any work experience in management). A possible explanation for that behavior is discussed below.

The simulator operates by participants developing and combining critical resources in appropriate levels in order to attract and retain customers, while running an efficient company. To reach this goal, the participants must understand the interdependence among critical resources and variables of the business system and they must combine these effectively. For instance, if they do not realize that Infrastructure, IT and Skilled Employees ought to be developed simultaneously so that appropriate values of Perceived Call Quality, IT Facilitation Index and HR Service Index are reached, they will not succeed in this game. In particular, the development of infrastructure (resource Base Stations) is critical to reach an acceptable level of customer satisfaction, and it has a significant influence on performance. Because suppliers take time to build capacity, this causes stock accumulation of BSIP and delays the process of

building infrastructure. The participants had to deal with such stock accumulation and delay effects. They had to anticipate the investment decision to access the required resource at the right time and to avoid stock accumulation in the later simulation periods (because such a situation implied investment in base stations that would not be operational during the simulation). MMS relies on concept relatedness and does not capture stock accumulations and delays comprehension. One explanation might be that participants with managerial experience deal better with dynamical effects, resulting in better performance. In order to explore this hypothesis, the variable BSIP Stock Accumulation is designed to measure the level of understanding of the accumulation and delay effects in the infrastructure building process.

Figure 12 shows a regression model that includes the variable BSIP Stock Accumulation. Regression analysis of BSIP Stock Accumulation on the most explanatory variables shows a strong effect of Work Experience ( $\beta = 0.418$ , p < 0.001) and MMS ( $\beta = 0.319$ , p < 0.01). Regression analysis of Performance on the most significant explanatory variables shows a strong effect of MMS ( $\beta =$ 0.438, p < 0.001), and BSIP Stock Accumulation ( $\beta = 0.286$ , p < 0.01). Since regression analysis on Performance as presented in Figure 11 showed a significant effect for Work Experience ( $\beta = 0.343$ , p < 0.01), and previous regressions indicated no significant effects for Work Experience, we also found a mediation effect of BSIP Stock Accumulation on the impact on Performance of Work Experience (Baron and Kenny, 1986). The participants' performance was seemingly influenced by the level of comprehension of the relatedness of critical variables (MMS) and recognition of the effects of stock accumulation and delay in building infrastructure (BSIP Stock Accumulation). This seems to



\**p* < 0.1; \*\**p* < 0.01; \*\*\**p* < 0.001

Fig. 12. Impact of BSIP stock accumulation: regression model with explanatory variables obtained through a stepwise regression (standardized betas). \*p < 0.1; \*\*p < 0.01; \*\*\*p < 0.001



Fig. 13. Historical behavior for participant X

have been influenced by managerial experience. Consequently, these results appear to confirm that participants with managerial experience tend to deal more effectively with stock accumulation and delay effects, and thereby achieve better levels of performance.

To illustrate some typical behavior, selected simulation results are presented in Figures 13, 14, and 16, showing the values of certain indicators for participant X from group B (high LBSC, low LSM), and Y and Z from group C (high LBSC, high LSM).

Figure 13 shows the historical behavior over time for participant X (who had no work experience), revealing poor Performance (14 M€) and low MMS (0.286). It can be observed that his firm kept increasing its level of investment on infrastructure as the BSIP and BS lines rose during the simulation, leading to a high BSIP value toward the end of the test. This suggests that this participant did not understand how stock accumulation and delay effects can impact infrastructure development (reducing investment in the final simulation periods would have increased performance). Because of his infrastructure development, Perceived Call Quality increased and reached a reasonable value at the end of simulation (0.99). The firm also invested intensely in IT and managed to maintain the IT Facilitation Index at its maximum value (1.2). However, this participant failed to understand the impact of HR Service Index and how to influence this variable. As the firm did not make correct decisions

in hiring and training employees, the customer satisfaction indicator progressively declined. Because of that, after period 11, Customer Satisfaction decreased to below a critical level and the firm lost all its customers.

As is evident from Figure 14, participant Y showed a different behavior. This participant (also without work experience) shows a very poor Performance (-432 M€) and low MMS (0.278). By looking at the BSIP and BS trends, one can conclude that the firm made strong infrastructure investments that became increasingly aggressive over time. Similar to X, these results indicate that Y did not understand the effects of stock accumulation and delays. Strong infrastructure development led to high values for Perceived Call Quality (1.37 at the end of the simulation). This participant also paid great attention to HR decisions as the HR Service Index shows reasonable values (1.22 at the end). However, this participant neglected the effects of the IT Facilitation Index. As can be seen in the final strategy map (Figure 15), this participant failed to relate IT Facilitation Index to Perceived Customer Service and, consequently, did not recognize the effect of IT Facilitation Index on Customer Satisfaction. At this point, it is interesting to note that Y also failed to draw correct arrows representing cause-effect relations. As the firm did not invest correctly in IT, the IT Facilitation Index indicator decreased progressively, reaching very low values during most of the simulation (0.36 by the end). However, the firm was able to remain in business because the level of Customer Satisfaction was



Fig. 14. Historical behavior for participant Y



Fig. 15. Final strategy map (causal diagram), drafted by participant Y of group C

reasonable, thanks to the very high values of Perceived Call Quality. Unfortunately, the enormous investment that resulted in such high Perceived Call Quality required a high capital expenditure, leading to very poor net performance.

Figure 16 shows the simulation results for participant Z (who had work experience), who obtained reasonable values for Performance (719 M€) and MMS (0.708). The strategy map drafted by Z is presented in Figure 10, and it shows that this subject understood the critical cause–effect relationships. Z invested strongly in infrastructure from the beginning until period 4. His investment was later discontinued, while the accumulated BSIP were progressively converted to active Base Stations, reaching and maintaining a reasonable resource level that ensured positive values for Perceived Call Quality. These results confirm that Z paid attention to the stock accumulation and delay effects in the process of building company infrastructure. Participant Z also managed to achieve and maintain acceptable values for the IT Facilitation Index and HR Service Index. Consequently, Customer Satisfaction was high



Fig. 16. Historical behavior for participant Z

and, combined with an adequate level of infrastructure investment, this led to good performance levels.

# Testing of hypotheses

Figure 17 shows the regression model including the most significant variables defined in the research model, showing a very significant impact of MMS on Performance ( $\beta = 0.538$ , p < 0.001). These results confirm two of the three hypotheses (Table 4): Hypothesis 2—The LSM positively influences MMS; and Hypothesis 3—MMS positively influences Performance.

On average, the participants from group B—BSC interface—showed a better MMS than subjects from group A— financial scorecard interface. These results are shown in Table 1, with a significant difference evident in Table 2. However, the regression showed no significant effect of LBSC on MMS. Consequently, our research does not provide full support of Hypothesis 1—that the LBSC positively influences MMS. The BSC performance measurement system provides better executive information than the financial scorecard, because the performance indicators of BSC capture the most relevant information about the system's behavior. Nevertheless, these results suggest that by using the BSC solely as a performance measurement system managers do not learn about the business system any more effectively than they would otherwise.





Table 4. Summary of hypothesis testing	Hypothesis	Description	Results
	H1	The Level of Balanced Scorecard positively	No significant
		influences Mental Model Similarity	correlation
	H2	The Level of Strategy Map positively	Supported
		influences Mental Model Similarity	
	H3	Mental Model Similarity positively	Supported
		influences Performance	* *

Using the strategy map for strategy review and implementation significantly improved the MMS of participants, supporting Hypothesis 2. Thus, the strategy map process seems to produce more effective double-loop learning. In fact, as we hypothesized, the results suggest that the process of strategy map review gave participants from group C a powerful systems thinking tool that accelerated their learning about the simulated business system.

We also found strong evidence of the fact that improved MMS led to better performance, in support of Hypothesis H3. Therefore, enhanced double-loop learning effectiveness (viewed as an improvement in mental models) seems to improve management performance.

The results also indicate the mediation effect in the BSC approach of Mental Model Similarity on performance. As shown in Table 5, LSM significantly influences MMS. The regression analysis "Performance (1)" shows an effect of LSM on Performance ( $\beta$ =0.248, p<0.1). When MMS is added to the regression analysis shown as "Performance (2)", MMS significantly influences Performance

Independent variables	Dependent variables							
	Mental Model Similarity		Performance (1)		Performance (2)			
	Standardized beta	Significance p	Standardized beta	Significance p	Standardized beta	Significance p		
LBSC LSM MMS	0.191* 0.529**	0.075 0.000	0.068 0.248*	0.608 0.066	-0.047 -0.073 0.606**	$0.695 \\ 0.595 \\ 0.000$		

Table 5. Regression analysis: test for mediation of Mental Model Similarity

p < 0.1; p < 0.001.

( $\beta = 0.606$ , p < 0.001), while the influence of LSM on Performance decreases greatly and is not significant ( $\beta = -0.073$ , p = 0.595). These results support the conclusion that MMS mediates the impact of LSM on Performance (Baron and Kenny, 1986).

#### Conclusion

This study is based on a laboratory experiment aimed at testing hypotheses about the impact of the BSC approach on students' and managers' learning and performance. Those hypotheses take into consideration a system dynamics and feedback learning perspective. The results of this simulation-based research provide useful contributions to the managerial field by showing how best to use the BSC approach in order to improve learning and performance. This study also offers contributions to the system dynamics field by reinforcing the importance of combining management flight simulators with causal diagramming modeling techniques as a basic strategy to improve mental models and dynamic decision making. This article also evidences the effectiveness and relevance of system dynamics modeling and simulation for experiments in the field of management research.

According to Kaplan and Norton, the BSC approach supports what Argyris calls "double-loop learning", facilitating and accelerating managers' understanding of cause-and-effect relations among essential components of the organization's strategy. The intermediate goal consists of seeding and improving managers' mental models about the business system, thereby leading to better performance. This research confirms the main assumptions already articulated in previous studies. We observed that the utilization of the BSC approach as a strategic management system (involving balanced scorecards and strategy maps as suggested by Kaplan and Norton) does improve managers' mental models of the structure of the business system. Results concerning the strong impact

of the causal diagram review process (strategy map) on learning and performance confirm that the feedback process for modeling and reviewing manager assumptions about cause-and-effect relationships leads to a better understanding of the business context and can promote organization performance. The participants' learning process in the simulated business context is only marginally improved when the strategy map is not reviewed. This is due to the fact that participants receive critical information but act as passive knowledge recipients as they internalize that information within their existing mental structures. The strategy map review strongly improves learning capacity because the participants become system modelers. Participants develop a systemic and dynamic understanding of the business context, by creating a causal model that represents the critical cause-and-effect relations. Our findings strongly confirm that the BSC approach can be significantly enhanced with the introduction of the strategy map concept. With this very simple systems thinking tool, the BSC approach offers managers opportunities to better model and learn about the business system.

This study also provides useful insights for accountants. We suggest that managers should complement the process of selection and utilization of strategic performance indicators with simple and effective modeling tools from the field of accounting. This process allows the review and improvement of managerial assumptions about the business context. Similar to other accounting studies (e.g., Ittner et al., 2003; Maiga and Jacobs, 2003; Hendricks et al., 2004; Braam and Nijssen, 2004), we found no evidence that the use of BSC as a performance measurement system enables managers to learn more effectively about the business system and improve its performance. Consistent with Lipe and Salterio (2000), this research also suggests that, due to cognitive limitations, managers may not appreciate the significance of non-financial and leading measures, thereby sidestepping the cause-and-effect logic chain and reducing the potential benefits of using the BSC. BSC usage only leads to higher performance if managers understand the cause-and-effect relations that link drivers with future financial performance. In other words, even though the set of performance indicators includes the most important components of the business system, this information is not helpful if managers do not understand the crucial relationships among these indicators and how they drive organizational performance. This conclusion confirms what Ittner and Larcker (2003) observed-namely, that many companies failed to use BSCs effectively because managers made little attempt to model and validate their understanding of the causal relationships between non-financial indicators and future financial performance. Finally, this paper provides important contributions to support the design of training programs that can teach and disseminate the BSC concept.

This study also provides a contribution to research that explores the performance implications of MMS. Our work reinforces the usefulness of the mental model construct as a means of investigating how managers learn about business systems and its impact on management performance in dynamical decision-making processes. The dynamic decision-making theory based on mental models asserts that managers make decisions which are the result of applying rules and policies governed by their mental models. An erroneous mental model means that there are significant differences between the managers' perception and the business reality. Several research questions regarding cognitive processes by which mental models are improved, and regarding the nature of how they determine intended actions and decisions, are the focus of other studies within the system dynamics community. One research question explores whether or not improved mental models of systems lead to better decisions. As suggested in similar previous research using the same simulator (Ritchie-Dunham, 2002), improved mental model similarity does lead to better performance. Another general research question asks which learning processes and strategies can best improve mental models. A significant number of promising techniques based on system dynamics modeling and simulation have emerged in order to improve mental models and dynamic decision making. Many experiments concerning the effects of interaction with management flight simulators have been carried out within the system dynamics community. This paper also documents some effects of interaction with system dynamics models on mental models. The results suggest that to improve mental models decision makers should take advantage of very simple systems thinking approaches like causal diagrams to model and review their understanding of the system. This research also emphasizes the importance of the "modeling for learning" concept. The system dynamics simulator used in this experiment appeared not to improve participants' learning when it was only applied as a management flight simulator without an accompanying modeling task. Causal diagramming processes led to new insights and reinforced the importance of modeling effectiveness when devising techniques and strategies for improving learning and dynamic decision making.

#### References

Akkermans H, van Oorschot K. 2005. Relevance assumed: a case study of balanced scorecard development using system dynamics. *Journal of the Operational Research Society* **56**: 931–941.

Argyris C. 1976. Increasing Leadership Effectiveness. Wiley: New York.

- Atkinson A, Balakrishnan R, Booth P, Cole J, Groot T, Malmi T, Roberts H, Uliana E, Wu A. 1997. New directions in management accounting research. *Journal of Management Accounting Research* **9**: 79–108.
- Baron R, Kenny D. 1986. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* **51**: 1173–1182.
- Braam G, Nijssen E. 2004. Performance effects of using the balanced scorecard: a note on the Dutch experience. *Long Range Planning* **37**: 335–349.

Camerer C, Hogarth R. 1999. The effects of financial incentives in experiments: a review

and capital–labor–production framework. *Journal of Risk and Uncertainty* **19**: 7–42.

Copeland T, Koller T, Murrin J. 2000. Valuation: Measuring and Management the Value of Companies. Wiley: Chichester.

Craik K. 1943. The Nature of Explanation. Cambridge University Press: Cambridge, UK.

- Davis F, Yi M. 2004. Improving computer skill training: behavior modeling, symbolic mental rehearsal, and the role of knowledge structures. *Journal of Applied Psychology* **89**: 509–523.
- Doyle J, Ford D. 1998. Mental models concepts for system dynamics research. *System Dynamics Review* **14**: 3–29.
- Doyle J, Ford D. 1999. Mental models concepts revisited: some clarifications and a reply to Lane. *System Dynamics Review* **15**: 411–415.
- Forrester J. 1961. Industrial Dynamics. Pegasus Communications: Waltham, MA.
- Gary M, Wood R. 2007. Testing the effects of a system dynamics decision aid on mental model accuracy and performance on dynamic decision making tasks. In *Proceedings of the 25th International Conference of the System Dynamics Society*, Boston, MA.
- Größler A. 2001. Musings about the effectiveness and evaluation of business simulators. In *Proceedings of the 19th International Conference of the System Dynamics Society*, Atlanta, GA. (CD-ROM).
- Hendricks K, Menor L, Wiedman C. 2004. The balanced scorecard: to adopt or not to adopt? *Ivey Business Journal* November/December: 1–7.
- Howie E, Sy S, Ford L, Vicente K. 2000. Human–computer interface design can reduce misperceptions of feedback. *System Dynamics Review* **16**: 151–171.
- Ittner C, Larcker D. 1998. Innovations in performance measurement: trends and research implications. *Journal of Management Accounting Research* **10**: 205–238.
- Ittner C, Larcker D. 2003. Coming up short on nonfinancial performance measurement. *Harvard Business Review* November: 88–95.
- Ittner C, Larcker D, Randall T. 2003. Performance implications of strategic performance measurement in financial services firms. *Accounting, Organizations and Society* **28**: 715–741.
- Johnson-Laird P. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness.* Harvard University Press: Cambridge, MA.
- Kaplan R, Norton D. 1992. The balanced scorecard: measures that drive performance. *Harvard Business Review* January–February: 71–79.
- Kaplan R, Norton D. 1996a. *The Balanced Scorecard*. Harvard Business School Press: Boston, MA.
- Kaplan R, Norton D. 1996b. Linking the balanced scorecard to strategy. *Californian* Management Review **39**: 53–79.
- Kaplan R, Norton D. 2000. Having trouble with your strategy? Then map it. *Harvard Business Review* September–October: 167–176.
- Kaplan R, Norton D. 2001a. *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment.* Harvard Business School Press: Boston, MA.
- Kaplan R, Norton D. 2001b. Transforming the balanced scorecard from performance measurement to strategic management. Part II. Accounting Horizons 15: 147–160.
- Kaplan R, Norton D. 2004. Strategy Maps: Converting Intangible Outcomes into Tangible Outcomes. Harvard Business School Press: Boston, MA.

- Lingle J, Schiemann W. 1996. From balanced scorecard to strategic gauges: is measurement worth it? *Management Review* **85**: 56–61.
- Lipe M, Salterio S. 2000. The balanced scorecard: judgmental effects of common and unique performance measures. *Accounting Review* **75**: 283–298.
- Maiga A, Jacobs F. 2003. Balanced scorecard, activity-based costing and company performance: an empirical analysis. *Journal of Managerial Issues* **15**: 283–301.
- Malmi T. 2001. Balanced scorecards in Finnish companies. *Management Accounting Research* **12**: 207–220.
- Mathieu J, Goodwin G, Heffner T, Salas E, Cannon-Bowers J. 2000. The influence of shared mental models on team process and performance. *Journal of Applied Psychology* **85**: 273–283.
- Mathieu J, Goodwin G, Heffner T, Salas E, Cannon-Bowers J. 2005. Scaling the quality of teammates' mental models: equifinality and normative comparisons. *Journal of Organizational Behavior* **26**: 37–56.
- Mooraj S, Oyon D, Hostettler D. 1999. The balanced scorecard: a necessary good or an unnecessary evil? *European Management Journal* **17**: 481–489.
- Olve N, Roy J, Wetter M. 2000. *Performance Drivers: A Practical Guide to Using the Balanced Scorecard*. Wiley: Chichester.
- Olve N, Petri C, Roy J, Roy S. 2003. *Making Scorecards Actionable: Balancing Strategy and Control.* Wiley: Chichester.
- Paich M, Sterman JD. 1993. Boom, bust, and failures to learn in experimental markets. *Management Science* **39**: 1439–1458.
- Repenning N. 2003. Selling system dynamics to (other) social scientists. *System Dynamics Review* **19**: 303–327.
- Richardson G, Andersen D, Maxwell T, Stewart T. 1994. Foundations of mental model research. In *Proceedings of the International System Dynamics Conference*, Stirling, UK.
- Rigby D. 2001. Management tools and techniques: a survey. *California Management Review* **43**: 139–160.
- Ritchie-Dunham J. 2001. Informing mental models for strategic decision making with ERPs and the balanced scorecard: a simulation-based experiment. In *Proceedings* of the 19th International Conference of the System Dynamics Society, Atlanta, GA. (CD-ROM).
- Ritchie-Dunham J. 2002. Balanced scorecards, mental models, and organizational performance: a simulation experiment. PhD dissertation. University of Texas at Austin, Austin, TX.
- Ritchie-Dunham J, Morrice D, Anderson E, Dyer J. 2007. A simulation exercise to illustrate the impact of an enterprise system on a service supply chain. *INFORMS Transactions on Education* **7**(3).
- Ritchie-Dunham J, Puente LM. 2008. Strategic clarity: actions for identifying and correcting gaps in mental models. *Long Range Planning* **41**: 509–529.
- Rowe A, Cooke N. 1995. Measuring mental models: choosing the right tools for the job. *Human Resource Development Quarterly* **6**: 243–255.
- Sengupta K, Abdel-Hamid T. 1993. Alternative conceptions of feedback in dynamic decision environments: an experimental investigation. *Management Science* **39**: 411–428.
- Schvaneveldt R. 1990. Pathfinder Associative Networks: Studies in Knowledge Organization. Ablex: Norwood, NJ.
- Senge P. 1990. The Fifth Discipline. Doubleday: New York.

- Speckbacher G, Bischof J, Pfeiffer T. 2003. A descriptive analysis on the implementation of balanced scorecards in German-speaking countries. *Management Accounting Research* 14: 361–387.
- Sterman JD. 1989. Modeling managerial behavior: misperceptions of feedback in a dynamic decision making experiment. *Management Science* **35**: 321–339.
- Sterman JD. 2000. Business Dynamics: Systems Thinking and Modeling for a Complex World. Irwin McGraw-Hill: New York.
- Stivers B, Covin T, Hall G, Smalt S. 1998. How nonfinancial performance measures are used. *Management Accounting* **44**: 46–49.
- Stout R, Salas E, Kraiger K. 1996. The role of trainee knowledge structures in aviation team environments. *International Journal of Aviation Psychology* 7: 235–350.
- Strohhecker J. 2007. Does a balanced scorecard management cockpit increase strategy implementation performance? In *Proceedings of the 25th International Conference of the System Dynamics Society*, Boston, MA.
- Warren K. 2002. Competitive Strategy Dynamics. Wiley: Chichester.
- Wyman B, Randel J. 1998. The relation of knowledge organization to performance of a complex cognitive task. *Applied Cognitive Psychology* **12**: 251–264.