Effects of time scale focus on system understanding in decision support systems

David N. Ford
Texas A&M University
Dalton E. M. McCormack
University of Bergen

Successfully managing dynamic complex systems requires an understanding of how structure influences both short- and long-term behavior. Therefore, decision support systems designed to improve performance by increasing user understanding require features that address both short and long time scales. The authors report the results of empirical research on the effects of features that facilitate different time scale focuses by users of management flight simulators on system understanding. System understanding was measured in two ways: with questions about structural relationships and by performance measures pertaining to the management of a complex system. Participants were divided into two time scale groups. Results were disaggregated based on causal distance and the timing of impacts to relate time scale focus and system understanding. A second experiment evaluated and improved the hypothesis to include the interaction of the time scale of system control and the time scale focus on improving system understanding and performance in managing dynamic systems.

KEYWORDS: decision support systems; time scales; management flight simulator; system dynamics; system understanding; learning.

Managing complex dynamic organizations and processes is critical to success in many systems such as national economies, business operations, and ecosystems. To control these systems managers must repeatedly make effective decisions, often with incomplete knowledge and information. The complexity of the causal structure and dynamic behavior of these systems disguise the impacts of specific managerial decisions to such an extent that knowing with certainty that decisions are satisfactory becomes difficult or impossible. Decision support systems (DSSs) are computerized aids designed to improve such decision-making processes (Singh, 1998). By providing managers with additional capacity to store, retrieve, and process information, DSSs expand the size and complexity of the systems that managers can perceive, understand, and control. Often, DSSs use simulation models to represent relevant portions of the system being managed, to store large quantities of information about the structure of the system, and to mimic the response of the actual system to its structure, environment, and managerial decisions. These simulation-based DSSs can be used as analytic tools for predicting the outcomes of decisions or as tools to expand the cognitive capacity of users.

AUTHORS' NOTE: The authors thank professor James Doyle for experiment design evaluation and the anonymous reviewer for comments.

© 2000 Sage Publications, Inc.
Management flight simulators (MFSs) are model-based tools that are often used as DSSs to improve decision making by facilitating user learning (Lane, 1995; Spector & Davidsen, 1997). When used as DSSs, these tools enhance learning by creating a safe artificial environment that reflects the important aspects of the system being managed and allows users to act as system managers within that environment (Bakken, Gould, & Kim, 1992). Consistent with other DSS researchers (e.g., O'Keefe, 1987; Warren, Van Dijk, & Johing, 1997), we equate learning in this context with improving the understanding of the system. In the case of managing a dynamic system this includes understanding the causal structure, behavior, managers and operators of the system, their policies, and the interaction of structure and behavior. For example, a sales manager who is unsure about the impact of a proposed pricing policy on company profit may implement the policy in an MFS that represents the response of the company, competitors, and customers to proposed pricing policies and use the MFS to experiment with different policy alternatives. The simulation compresses the time required for the sales manager to see system responses that appear in the real system over long periods of time to just seconds. This provides the sales manager an opportunity to learn how the structures modeled in the MFS and his pricing policy interact to drive profits. As a safe managerial practice field MFSs can improve managerial decision-making performance by improving managers' understanding of the interconnected nature of systems and the consequences of their actions (Senge, Roberts, Ross, Smith, & Kleiner, 1994).

The effects of differences in a system's responses in different time scales on managerial performance illustrate how MFSs can improve decision making by facilitating learning. The delayed responses of dynamically complex systems to managerial decisions and actions can be significantly different than short-term responses. For example, a large, sudden price increase may raise profit in the first few months following the increase but also drive away customers and decrease profits in future years. Large dynamic systems are, therefore, often perceived and managed on several time scales. For example, in fighting a forest fire, the leaders of individual fire-fighting units make decisions about their units' locations and movements in response to local fire conditions in a time scale measured in hours. Team leaders coordinate individual fire-fighting units on a daily basis as fire lines move through the forest and the fire chief directs multiple teams and equipment to different areas of the forest over several weeks. In this example, each manager primarily sees and manages in a single time scale (hours, days, or weeks). But many systems and management structures require one manager to perceive, understand, and manage multiple time scales simultaneously. Brehmer (1998) describes one such management structure. This challenge raises several important MFS design questions. How do expert managers perceive different system responses in the multiple time scales that are inherent in dynamic systems? How do managers learn about managing different time scales? How can MFSs help managers learn about system responses in different time scales and the structures that cause them?

Understanding how time scales affect learning is important to MFS researchers and developers because understanding in multiple time scales is critical to successful management, and time scale features can influence the effectiveness of MFSs in facilitating user learning. A clear understanding of the relationships between time scale focus
and system understanding will give MFS designers guidelines for decisions such as the number of time scales they should include in MFSs to achieve maximum system understanding and performance. Accordingly, this article addresses the issue of how the time scale focus of managers affects learning.

In the next section, the specific problem of the relationship between time scale focus and learning is described and clarified. Then, we present our hypotheses and research method. The results of our experiment and our analysis are followed by a description of how we improved our conceptual hypothesis using a second experiment. We conclude with our assessment of the contributions to DSS research and design and suggestions for future research.

**Problem Description**

Designing MFSs for learning is difficult due to the lack of a model of learning about complex systems to guide MFS design. This model is not available partially because the necessary ingredients for learning about complex systems are not clearly understood. Instructional design approaches to designing DSSs provide general requirements (e.g., a learning theory and user-friendly interfaces) but are not specific enough to guide MFS design (Spector & Davidsen, 1997). According to Sterman (1994), such a model also must address the causes of failures to learn: dynamic complexity, imperfect information about the state of the real world, confounding and ambiguous variables, poor scientific reasoning skills, defensive routines and other barriers to effective group processes, implementation failure, and misperception of feedback. Increased experience in managing these systems is a logical approach to improving understanding and thereby improving managerial performance. However, empirical researchers have shown that little learning occurs when decision makers are asked to manage dynamic complex systems (Paich & Sterman, 1993), due largely to failures to understand the delayed impacts of system structure on behavior (Diehl & Sterman, 1995). Therefore, designing MFSs to provide opportunities for managers to practice decision making is not sufficient for learning.

Despite the lack of an underlying model for MFS design, researchers have suggested general design guidelines. For example, Senge and Sterman (1994) suggest that the design of effective learning labs (of which MFSs form a central part) should focus on conceptualization and design opportunities for reflection and should not focus learning on the computer. However, more specific guidance is needed to provide sufficient knowledge and processes for learning. For example, even if users are provided a complete description of the system model and all the relationships between variables, they will find it difficult to successfully manage the system (Sterman, 1994). This is because understanding the relationships between variables reflects a comprehension of only the structure of the system and not the relationship of structure and behavior, which is necessary for learning and effective management (Forrester, 1961). Sterman (1994) argues that features for developing an understanding of structure-behavior interactions are necessary components of MFSs designed for learning, but he does not
provide a framework for identifying the sufficient set of characteristics, conditions, or features. Kahneman and Tversky describe characteristics of information feedback often provided to users that prevent learning:

Effective learning takes place only under certain conditions: it requires accurate and immediate feedback about the relation between the situational conditions and the appropriate response. The necessary feedback is often lacking for the decisions faced by managers, entrepreneurs, and politicians because (i) outcomes are commonly delayed and not attributable to a particular action; (ii) variability in the environment degrades the reliability of the feedback, especially where outcomes of low probability are involved; (iii) there is often no information about what the outcome would have been if another decision had been taken; (iv) most important decisions are unique and therefore provide little opportunity for learning. (as cited in Bakken, 1993, p. 25)

Although Kahneman and Tversky's criteria for feedback do not necessarily lead to optimal learning under all conditions, they can be useful for improving learning in complex systems. Many MFSs (including the one described here) meet most of Kahneman and Tversky's guidelines. Paich and Sterman (1993) recommend that MFS designers provide different forms of information feedback based on their cognitive impact on users, such as outcome feedback to allow comparison of performance with goals and cognitive feedback to improve their structural models of the system. However, research in the roles of different types of feedback on learning about dynamic complex systems has not produced useful design guidelines. Few guidelines are available for the effective design of MFSs that identify design features or other practical guides for MFS development. Those that are available are directed almost solely toward the improvement of managerial performance and not learning per se (e.g., Langley & Morecroft, 1996; Machuca, Castillo, Carrillo, & Zamora, 1998; Sengupta & Abdel-Hamid, 1993).

Research on learning with MFS design has also focused on improving managerial performance (Bakken, 1993; Dörner, 1989; Ford, 1999; Langley, 1996; Sengupta & Abdel-Hamid, 1993). Many researchers have used improvement in managerial performance as a measure of learning. However, Bakken (1993) showed that learning is not always correlated with improvement in decision-making performance by measuring learning in two different forms—improvement in dynamic decision-making performance and improvement in causal understanding—and finding no correlation. Similarly, Warren et al. (1997) found that the availability of a simulation model in a DSS improved performance concerning the dynamics of the system but had no positive impact on system understanding. Because performance is a poor indicator of learning, research directed solely toward performance provides little guidance in MFS design for learning. A broader measure of learning is needed to evaluate MFS effectiveness for this purpose.

Researchers have found it difficult to measure learning in dynamic environments. Some systems dynamics researchers have used changes in mental models as a proxy for learning (Bakken, 1993; Doyle, Radzicki, & Trees, 1996). This is problematic because the meaning of the term "mental models" is not clearly established in system
dynamics. According to Doyle and Ford (1998), although "mental models are of central importance to system dynamics research and practice, the field has yet to develop an unambiguous and agreed upon definition of them" (p. 3). Some of these problems are highlighted in an experiment by Doyle et al. (1996) in which learning was measured as changes in mental models. Changes in mental models were defined as changes in understanding of causal relationships in the system. They found evidence of a positive relationship between MFS use and changes in mental models. However, the researchers found the data extremely difficult to interpret because mental models change quickly and deriving methods to collect and measure these changes is difficult. This research measured learning solely with changes in mental models and not with dynamic decision-making performance. This prevented an assessment of whether participant mental model changes resulted in learning or a loss of system understanding. This research shows that MFSs can change mental models and can potentially improve learning. But to reliably test MFSs for learning requires measures of both mental model change and the results of those changes, that is, both causal understanding and managerial performance.

Despite the lack of specific design guidance, the need to improve user understanding in multiple time scales is considered critical for learning about complex systems. We conceptually define a managerial time scale as the time between system evaluations and assessments by a manager relative to the period of the unmanaged system. Therefore, managers using a short managerial time scale evaluate and assess the system frequently compared to managers using a long time scale. These refocusing events are most likely to occur when decisions are required or changes in decisions are possible. Paich (1994) and Sterman (1989) have shown that when managers are asked to make frequent decisions in dynamic environments, their decisions usually involve very simple rules. These heuristics are constructed to solve short-term problems and fail to recognize delays, phase shifts, and other important structural relationships and how they affect system behavior. Paich and Sterman (1993) vividly document the short managerial perspective of their subjects when making frequent management decisions with logs of subject strategies throughout MFS use. According to Granger, the reason for the adoption of a short time scale focus by managers is that "although a variable may affect a dependent variable, if the effect is subsumed by the past data series, the causality cannot be explicitly grasped" (cited in Lee & Yum, 1998, p. 137). This misperception of delayed cause-effect relationships limits managerial performance (Diehl & Sterman, 1995). Brehmer (1998) suggests that the problem may not be that managers are incapable of learning about dynamic complex systems or managing them but that they cannot do both simultaneously. Therefore, the use of simple rules and focusing on one time scale at a time may be strategies for learning about and managing complex systems within the limits of bounded rationality. Bounded rationality describes the limitations on human cognitive processing as central to understanding human behavior and performance (Simon, 1974, 1996). Regardless of the strategies used in practice, successfully addressing multiple time scales is critical to dynamic system management and therefore an essential part of MFS design. This makes explicitly incorporating time scale effects into MFS designs critical.
The design of effective time scale features in MFSs requires an understanding of how the length of users' time scale focus influences changes in system understanding. But the relationship between time scale features and system understanding improvement is not understood adequately to reliably design effective MFSs for learning. An improved measure of system understanding improvement (learning) in dynamic systems is required to develop useful insights into the nature of this relationship. This research investigates the relationship between time scale focus length and improvement in system understanding in MFSs using multiple measures of learning.

**Hypotheses**

We conceptually hypothesize that longer time scale focuses facilitate learning more than shorter time scale focuses. Systems theory suggests that managers need a perspective that is long enough to allow them to see and understand a system's delayed responses as well as its immediate responses to develop an accurate understanding of the system (Forrester, 1961). More specifically to MFSs, Simons (1990) and Langley and Morecroft (1996) have suggested that time scales that are relatively long facilitate learning more than do shorter time scales. According to Langley and Morecroft (1996),

"Switching the mode of the user interaction with the microworld [MFS] from gaming [shorter time scale] to simulation [longer time scale], allowing users to specify policies rather than decisions and run the model for continuous periods may improve their ability to identify high leverage policies. It is likely to allow more efficient use of time available, and an opportunity to apply the scientific method to systematically investigate the policy space. (p. 303)"

They reason that managers using longer time scales will have more time to design and test their strategies as well as to reflect on the behavior that these strategies generate. Langley and Morecroft (1996) also suggest that managers in shorter time scales usually change their strategies during simulations, disrupting the system's response to earlier strategies and unintentionally obfuscating the system's long-term response. This causes short time scale users to have a limited, different, and inferior understanding of the system compared to users with longer perspectives. These effects would be expected to lead to poorer managerial performance by users with shorter time scale focuses. This is supported by Dörner's (1989) finding that reflection improved simple decision-making performance.

We operationalize our hypothesis as follows. The length of managerial time scale focus is our independent concept, which we quantify in two lengths: long and short time scales. Directly measuring lengths of participant perspective is difficult. Therefore, we measure time scales with the decision time interval divided by the period of the unmanaged system so that longer time scales are represented with larger numbers and to normalize across systems. As described above, we equate learning, our dependent concept, with improvement in system understanding. We disaggregate system understanding improvement into two concepts—changes in mental models and managerial
performance—which we measure independently. System dynamics research suggests a link between changes in mental models and improved system understanding (Doyle et al., 1996; Senge et al., 1994). We measure changes in mental models with changes in the quantity and direction of managerial understanding of the causal relationships between pairs of system model variables (causal understanding improvement). This is the same approach and measurement used by Doyle et al. (1996) and Bakken (1993) to indicate changes in mental models. Managerial performance is used to indicate the direction and size of mental model change impacts and to relate treatment effects to the ultimate objective of DSSs (performance enhancement). We measure managerial performance with the cumulative deviation of the managed system behavior from a target behavior. Based on our conceptual hypothesis and operationalization, a change in mental models indicating more understanding of causal relations reflects learning and better managerial performance supports a finding of learning. Therefore, we propose and test two operational hypotheses.

Hypothesis 1: Managers using a long time scale focus will increase their causal understanding while using a MFS more than managers using a short time scale focus.

Hypothesis 2: Managers using a long time scale focus will have less cumulative deviation of system behavior from a target while using a MFS than managers using a short time scale focus.

Several factors could intervene in the effect of time scales appearing in our results. Primary among these confounding factors is the complexity of the system being managed. We suspect that the bounded rationality of participants is often exceeded by the complexity of MFSs. Significantly exceeding participants' bounded rationality may cause the loss of treatment effects due to the overwhelming difficulty of the task. Other potentially intervening factors include manager training and experience in systems, management techniques, and domain knowledge specific to the system.

Research Methodology

We used a controlled experiment to test our hypotheses. Participants used a custom-built MFS to play the role of a manager of an isolated ecosystem attempting to control the size of a deer herd through hunting, importing predators, and planting grass. Details of the research design, including a description of the computerized questionnaire, interface, measured variables, and decision-making task, are presented in the following sections. A description of how the experimental sessions were designed and conducted, and how research participants were secured, is provided in the following sections. McCormack (1998) provides additional details.

Computerized Causal Questionnaire

Research participants used a questionnaire developed with a customized Delphi database management application program to describe their understanding of the
hidden simulation model of the ecosystem. The multiple choice questionnaire contained 25 questions on causal relationships between variables in the model. The form of these questions (see Figure 1) was based on a similar experiment by Bakken (1993). The five response choices were based on Huff, Narapareddy, and Fletcher's (1990) adaptation of Axelrod's (1976) method of coding causal relationships described in free form by participants and therefore are capable of capturing user understanding of causal relations. For example, the correct answer to the question, “An increase in grass eaten by deer leads to... in grass” (see Figure 1) is “delayed decrease” because an increase in “grass eaten by deer” will lead to a delayed decrease in the amount of grass available. Each question appeared on the monitor for 25 seconds regardless of whether the question was answered before it was replaced by the next question. An animated timer (see Figure 1, right side) kept participants informed of the time remaining before a new question appeared. Research participants were unable to return to previous questions. Each participant saw all 25 questions. Answers were stored in a database for analysis. To control for sequence effect and to keep participants interested and motivated to perform well on the posttest questionnaire, the order of the questions was randomized across tests.

**MFS Interface**

The MFS interface (see Figure 2) contained a numerical management performance indicator, graphical and tabular displays of system behavior, and input features for decision parameters. The interface was developed as a POWERSIM application that allowed participants to use the underlying model of the ecosystem to simulate the behavior of the deer population in response to the ecosystem structure and the values of the four decision variables. Decisions were automatically captured by the system for analysis. This form of data collection was selected instead of verbal protocols to
FIGURE 2: The Kaibab Plateau Management Flight Simulator Interface

prevent interference with decision making by forcing individuals to verbalize their thought processes and actions (Cook & Swain, 1993).

Experimental Measures

Effectively capturing treatment effects required measures of causal understanding and managerial performance. Data collection included the use of the causal understanding questionnaire, pre- and postexperiment qualitative data, and managerial performance. Questions on the causal understanding questionnaire were based on the causal-loop diagram (Goodman, 1974; Richardson & Pugh, 1981) of the ecosystem model, which describes the most important relations among system components (see Figure 3). Participant performance on the causal questionnaire was measured according to a scoring system used by Bakken (1993) in which one half point was awarded for correctly identifying the direction of causal influence (increase or decrease), and an additional half point was awarded for correctly identifying the timing of the influence (immediate or delayed) if the direction was identified correctly. The qualitative data collected included questionnaires concerning comprehension of the passage describing the ecosystem, ease of causal questionnaire use, strategies used to manage the system, and task complexity. Control data also were also collected for participant age, gender, primary language, English proficiency, and five types of education and work experience.
FIGURE 3: Causal Loop Diagram of Kaibab Plateau Simulation Model
NOTE: -> = a causal relationship: change in variable at arrow's tail causes change in variable at arrow's head; + = variable at arrow's head moves in the same direction as the variable at the arrow's tail; - = variable at arrow’s head moves in the opposite direction as the variable at the arrow’s tail, it = delay in response of variable at arrow’s head to change in variable at arrow's tail. R = reinforcing feedback loop: in isolation generates exponential growth or decay; and B = balancing feedback loop: in isolation generates goal-seeking behavior. Feedback loops are as follows: B1 = predators eat deer and control both populations, B2 = deer eat grass and control both deer and grass, B3 = deer density controls deer eaten, B4 = grass availability controls grass eaten, B5 = deer control grass availability and deer health, R1 = predator population growth or decay, R2 = deer population growth or decay, and R3 = grass growth.

Managerial performance was measured by how close participants kept the deer population to the target population of 30,000. Performance was quantified with the cumulative variance over the 40-year time horizon managed by the participants. Participants were able to monitor this figure on the MFS interface.

Research Design

Several experimental designs were considered, including combinations of different time scale focus treatments applied in various sequences. Based on the early development of this line of research, the strength of the experimental design, and the limited number of available participants, a design was selected that compared the impacts of two time scale focus lengths on improvement in causal understanding and managerial performance (see Table 1). We describe our use of an alternative experiment design later.
### TABLE 1: Experimental Design and Parameter Measures

<table>
<thead>
<tr>
<th>Group</th>
<th>Pretest</th>
<th>Treatment</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long time scale</td>
<td>Causal questionnaire</td>
<td>Long time scale tool</td>
<td>Causal questionnaire and performance</td>
</tr>
<tr>
<td>Short time scale</td>
<td>Causal questionnaire</td>
<td>Short time scale tool</td>
<td>Causal questionnaire and performance</td>
</tr>
</tbody>
</table>

The difference between the two treatment groups was the simulated time between management decisions. Participants in the short time scale group selected decision variable values annually throughout the 40-year managed time horizon. Participants in the long time scale group could select only one set of decision variable values and were unable to stop the simulation. Therefore, long time scale participants had only one opportunity to perceive the system, whereas participants in the short time scale group were able to change their decisions and revise their perception of the system 40 times. Time scale length was measured with the number of decisions made during a simulation (1 or 40) and described as the decision interval per unmanaged system period (30 years). In a second experiment, other time scales were considered, as discussed in the Hypothesis Evaluation and Improvement section.

### The Managerial Task

The decision-making task consisted of setting the values of four decision variables (grass seeding, deer hunting, predator importation, and predator hunting) to control the deer population. We assume that participants sought an improved understanding of the system (i.e., learning) to attain this goal, which is supported by our postexperiment interviews with participants. A simplified ecological system was chosen to partially control for participant-bounded rationality. Bakken (1993) has shown that domain familiarity can influence causal understanding and performance. Therefore, we designed the descriptions of the system and task, the underlying model, and the interface to create a relatively simple artificial environment with low domain knowledge requirements. The simulated ecosystem is based on a system dynamics model of the Kaibab Plateau (Goodman, 1974; Roberts, Andersen, Deal, Garet, & Shaffer, 1983; Sterman, n.d.). Our specific version of this model has been previously used to study how information structures affect performance in a policy-development environment (Ford, 1997). Ford’s version of this predatory-prey model simulates the interaction of three species (dear predators, deer, and grass) over time. Despite its structural simplicity (24 equations), the system remains dynamically complex. Documented model equations are available from the authors.

### Research Participants and Experimental Protocol

Our target population is practicing managers of dynamic systems. Due to the unavailability of managers, undergraduate and graduate college students were used as
subjects. The use of students as surrogates for practicing managers is considered adequate because this is an information-processing task in which students and managers are expected to behave in similar ways (Singh, 1998). Thirty-nine individuals participated. One participant’s data were discarded because of problems the individual encountered with the experimental tools. Thus, 38 data points were collected (19 in each time scale group). The experiment was run at two colleges to obtain access to additional participants and to control for educational and environmental biases. Participation in the study was voluntary; however, incentives were used. Participants in Bergen, Norway, were paid 100 Norwegian Kroner (about U.S.$13) for participating. An additional 100 Kroner were awarded to the participant with the highest average score in the causal questionnaire in each group. Participants in Worcester, Massachusetts, were awarded U.S.$50, U.S.$20, and U.S.$10 for the best, second, and third score based on their causal understanding improvement. The difference in performance improvement between the two participant groups (i.e., Bergen and Worcester) was not significant. Therefore, differences in incentives between the two locations are not believed to have affected results. Individuals volunteered for a particular session on the basis of availability, with each session lasting a maximum of 2 hours.

Each session began with a description of the Kaibab Plateau (based on Goodman, 1974) and the causal understanding questionnaire. The pretest questionnaire was exactly the same for each treatment group. Participants in both groups then used the MFS for 20 minutes to stabilize the deer population at a level of 30,000 for a period of 40 years (1940-1980). In each simulation, participants clicked on the “run” button on the command bar of the MFS interface. The simulation starts in the year 1900 and stops at the year 1940. The ecosystem description and its cyclical behavior from 1900 through 1940 provided all participants with a common historical behavior on which to base their management of the ecosystem and the information needed to reflect on the system before making initial managerial decisions if they chose to do so. At this point, participants typed their decision variable values into the input boxes and implemented their decisions until the simulated year reached 1980. Each simulation of the ecosystem over the 80-year time horizon was saved on a diskette. At the end of the 20-minute session, all participants completed a posttest questionnaire. The posttest questionnaire was the same as the pretest questionnaire except that the order of the questions was changed at random. After the posttest questionnaire, participants completed a different questionnaire to gather experimental control data and qualitative data as described previously.

Results

The presentation of the major results is divided into two sections. The first section presents a statistical testing of the research hypotheses. The second section discusses results of exploratory analyses performed to better understand the results.
TABLE 2: Causal Understanding and Managerial Performance Means and Causal Understanding Improvement

<table>
<thead>
<tr>
<th>Group</th>
<th>Pretreatment Causal Understanding</th>
<th>Posttreatment Causal Understanding</th>
<th>Causal Understanding Improvement</th>
<th>Best Managerial Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long time scale focus</td>
<td>14.07</td>
<td>14.68</td>
<td>+0.61</td>
<td>2.16</td>
</tr>
<tr>
<td>Short time scale focus</td>
<td>13.97</td>
<td>14.55</td>
<td>+0.58</td>
<td>2.65</td>
</tr>
<tr>
<td>Treatment effect significance</td>
<td></td>
<td></td>
<td>*p &gt; .10 (ns)</td>
<td>*p &gt; .10 (ns)</td>
</tr>
</tbody>
</table>

Statistical Analyses

For each group, the causal understanding scores of the pre- and posttest questionnaires and causal understanding change were computed. Pretreatment scores for the two groups are not significantly different, supporting our use of causal understanding improvement as a measure of learning and our assumption of negligible individual participant differences. The best system performance of both groups also was computed. One-sided t tests for statistical significance of group differences in causal understanding improvement and managerial performance were performed. Table 2 shows the pretreatment and posttreatment causal understanding questionnaire scores and causal understanding improvement of both groups. The causal understanding of both groups increased but the difference was not significant (*p > .10*). However, the impact of the time scale on causal understanding improvement is larger for the long time scale than the short time scale, as expected. A comparison of the best managerial performance for each group shows that the short time scale focus group performed better than the long time scale focus, although the difference is not statistically significant. No significant correlations were found between the intervening variables and causal understanding or managerial performance.

Based on the lack of a significant positive difference between the causal understanding improvement of the long time scale group over the short time scale group, our results do not support Hypothesis 1. Based on the lack of a significant positive difference between the managerial performance of the long time scale group compared to the short time scale group, our results also do not support Hypothesis 2. We found no correlation between managerial performance and causal understanding, supporting a similar finding by Bakken (1993). This result contrasts sharply with the intuitive belief that improved causal understanding improves performance. The absence of necessary requirements for improved performance in our experiment can explain these results.

Of interest, the expected larger understanding improvement in the long time scale group contrasts with the group’s unexpected lower managerial performance. This suggests that a deeper analysis may reveal additional important factors in the relationship between time scales and system understanding. Several characteristics of dynamic systems and how managers attempt to control them could contribute to this relation-
TABLE 3: Number of Causal Understanding Questions by Causal Distance and Timing

<table>
<thead>
<tr>
<th>Causal Distance</th>
<th>Immediate Impact</th>
<th>Delayed Impact</th>
<th>All Causal Relationships</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close</td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Medium</td>
<td>1</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Distant</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>All Causal Relationships</td>
<td>9</td>
<td>16</td>
<td>25</td>
</tr>
</tbody>
</table>

ship, including the closeness of causally related components, the timing of impacts, and the flexibility of user control of the system. Next, we describe how we used additional statistical analysis of the data collected in the experiment described above to investigate the role of causal closeness and timing of impacts. This is followed by additional hypothesis development centered around the flexibility of managerial control and based on a second experiment.

Exploratory Analyses

Disaggregation of causal understanding data. We suspect that the impacts of time scale focus on managers are different for different types of causal relations within a system. More specifically, long time scales may improve causal understanding more for causal relations that are relatively short and are therefore easier to understand than for causal relations that are related through many intermediate causal relations, which make them more difficult to understand. In addition, we suspect that long time scales may improve causal understanding more for causal relations that have immediate impacts and are therefore easier to understand than for causal relations that have delayed impacts, which make them more difficult to understand. To investigate this concept further and to better understand how time scale focus length impacts system understanding improvement, we disaggregated the causal understanding questionnaire data in two dimensions: causal distance and timing (see Table 3). There were three causal distances: close, medium, and distant. The nine causally close questions tested participant understanding of the relationships between two variables, which were separated by one or two causal links as measured on the causal loop diagram of the ecosystem model (see Figure 3). The 12 medium questions involved variables separated by three to five causal links and the four causally distant questions tested the understanding of variables with six or more causal links between them. We also disaggregated the causal understanding questions between those in which the first variable has an immediate impact on the second and those with delayed impacts. Although the option "no change" was given as an answer alternative in the questionnaire interface, no questions had this as the correct answer. Therefore, all 25 questions are included in our $2 \times 3$ matrix. There is either 0 or 1 question in the immediate/medium and immediate/distant cells of Table 3 because causal relations that relate variables located relatively far apart are delayed in the model and in most actual
TABLE 4: Additional Hypotheses Based on Disaggregated Data

<table>
<thead>
<tr>
<th>Causal Distance</th>
<th>Immediate Impact</th>
<th>Delayed Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close</td>
<td>H3</td>
<td>H4</td>
</tr>
<tr>
<td>Medium</td>
<td>NH</td>
<td>H5</td>
</tr>
<tr>
<td>Distant</td>
<td>NH</td>
<td>H6</td>
</tr>
</tbody>
</table>

NOTE: H3 = Hypothesis 3, H4 = Hypothesis 4, H5 = Hypothesis 5, H6 = Hypothesis 6, NH = no hypothesis tested due to insufficient number of questions for analysis; see Table 3.

TABLE 5: Differences Between Long and Short Focus Causal Understanding Improvement for Additional Hypothesis Tests

<table>
<thead>
<tr>
<th>Causal Distance</th>
<th>Immediate Impact</th>
<th>Delayed Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close</td>
<td>(+0.60) p &gt; .10</td>
<td>(-0.16) p &gt; .10</td>
</tr>
<tr>
<td>Medium</td>
<td>NH</td>
<td>(-0.29) p &gt; .10</td>
</tr>
<tr>
<td>Distant</td>
<td>NH</td>
<td>(+0.34) p &gt; .10</td>
</tr>
</tbody>
</table>

NOTE: NH = no hypothesis tested due to insufficient number of questions for analysis; see Table 3.

systems. We consider these two cells to provide insufficient data for useful hypothesis testing.

Additional hypotheses. Based on this disaggregation, we tested the following additional hypotheses (see Table 4):

Hypothesis 3: Managers using a long time scale focus will increase their causal understanding of immediate causal links separated by small causal distances while using a MFS more than will managers using a short time scale focus.

Hypothesis 4: Managers using a long time scale focus will increase their causal understanding of delayed causal links separated by small causal distances while using a MFS more than will managers using a short time scale focus.

Hypothesis 5: Managers using a long time scale focus will increase their causal understanding of delayed causal links separated by medium causal distances while using a MFS more than will managers using a short time scale focus.

Hypothesis 6: Managers using a long time scale focus will increase their causal understanding of delayed causal links separated by large causal distances while using a MFS more than will managers using a short time scale focus.

Table 5 summarizes the results of our analyses. No significant correlations were found. We suspect this is due to the relatively few questions used to test any one hypothesis. Therefore, we evaluate the data in a manner similar to Brehmer (1998) using the consistency of changes and direction of differences in data. The results are mixed in their support of the hypotheses, with two relations (close immediate and distant delayed) showing more improvement by long time scale participants than short time scale participants and the other two relations (close delayed and medium delayed) indicating less improvement by long time scale participants. A review of the pairs of
causal understanding improvement scores (for the long and short time scale groups) for each hypothesis reveals that the delayed medium (Hypothesis 5) and delayed distant (Hypothesis 6) pairs improved their scores and the causally close pairs (Hypotheses 3 and 4) reduced their scores. This suggests that the treatment helped participants learn about relations related to dynamic system features (those delayed and causally distant) but degraded system understanding of more static relations. We explain this result in the next section.

The results of our analyses of the disaggregated data are inconsistent in their support of the concept of a difference in time scale focus impacts based on causal distance or timing of impacts. Several experiment features can explain the apparently weak treatment effect. Despite our efforts to limit the complexity of the tasks, our system and questions may have exceeded the bounded rationality of our participants so far that treatment effects were lost in their confusion or the challenge of the decision-making task. An experiment in which participants manage an even simpler system can test this explanation. Our incentive to improve system understanding may have been inadequate to convince participants to seek understanding as well as system performance, resulting in a poor measure of learning. Despite our improvements, our measurement tool may inadequately capture causal structure understanding (e.g., feedback loop understanding is not captured). Inadequate treatment time (20 minutes) may have been provided to improve participant understanding. This is supported by Warren et al.'s (1997) finding that two 90-minute interactions with a simulation-based DSS within a 1-day workshop were adequate to improve performance but not system understanding in information systems professionals. The absence of conditions or features necessary for learning in dynamically complex systems also could explain our results. For example, Sterman (1994) suggests that bringing participants closer to the modeling process can improve learning in dynamic environments. Therefore, it is possible that a more sophisticated MFS in which participants can build feedback loops into the control of the system such as was used by Ford (1997) can generate larger improvements in causal understanding. If time scale focus impacts vary with causal distance or the timing of impacts, the effects may be secondary to other more influential but still incompletely identified factors that explain how time scales and system understanding are related.

Hypothesis Evaluation and Improvement

Our results generated more questions than answers. Is the impact of time scale focus length on system understanding improvement linear, truncated due to other effects, or very nonlinear? Why do the impacts on causal understanding improvement and managerial performance differ? What types of understanding and knowledge are required to improve system understanding? How can they be developed with MFSs? What measurement tools can capture different forms of understanding and learning?

To improve our hypotheses about how time scales affect learning through MFSs, we conducted a second experiment that addressed several of the experimental design
issues above. In particular, we used four time scales instead of two to allow a more specific description and analysis of the time scale-learning relationship and rewarded a balanced performance in system management and causal understanding to motivate participants to learn. McCormack (1998) describes this experiment and its results. Although our second experiment produced no statistically significant results, it proved useful for hypothesis development. As shown in Figure 4, participants' best managerial performance decreased as the time scale length increased. These results are consistent with the results of our first experiment but contrast with our original hypothesis that causal understanding improvement and performance would increase as the time scale length increases.

In contrast to the managerial performance data, causal understanding improvement in relation to time scale focus length in the second experiment peaks in the second shortest time scale (see Figure 5). A peak in causal understanding improvement occurs consistently in the second or third time scale when these results are disaggregated by causal distance (close, medium, or distant) and the timing of impacts (immediate or delayed), as described previously.

The difference in the shapes of Figures 4 and 5 suggests that our treatment affects managerial performance and causal understanding improvement differently. We suspect that our treatment influenced a second critical factor that affects learning differently than the length of the participant's perspective: the time scale of the participant's control of the system. Conceptually, the time scale of control is the relative time between opportunities for managers to change their decisions and alter the behavior of the system. The time scale of control differs from the time scale focus in reflecting the manager's ability to change system behavior instead of the manager's perception of the system. Control is directed toward performance, whereas focus or perspective is directed toward understanding. Control time scales can be measured with the time between managerial decisions in a manner similar to our measurement of time scale focus length. Managers need to have some form of control of the system to affect
performance. Longer time scales make control more difficult because managers have fewer chances to adjust the system (Hollnagel, 1998). This is consistent with our interviews with participants using longer time scales who reported frustration due to a lack of control. Very long control time scales would therefore be expected to reduce performance (see the right side of Figure 4). In contrast, participants using shorter time scales are able to make the necessary adjustments to improve performance (see the left side of Figure 4). A causal relationship between control time scales and managerial performance can explain our observed performance data better than our original hypothesis.

Control time scales can also help explain the peak in causal understanding improvement in intermediate time scale groups. To learn experimentally, participants must do at least two things: (a) perceive and analyze delayed system responses, reflect on their past decisions, and rationalize about what led to the behavior; and (b) test possible structures by observing system responses to changes. We suspect that long-time scale perspectives increase participants' ability to perceive, analyze, reflect, and rationalize, especially on the dynamic aspects of systems. A larger positive learning effect on understanding the critical dynamics of our system than on understanding static system characteristics is supported by learning in delayed/medium and delayed/distinct links described previously (see Table 5) and the increase in causal loop understanding that we found with longer time scales in our second experiment (see McCormack, 1998, for details). Understanding causal loops reflects a comprehension of dynamics relative to understanding causal links that reflect static system relations. Warren et al.'s (1997) finding that dynamic simulation improved participant understanding of dynamic system features but not static system features also supports this hypothesis. In contrast, we suspect that longer control time scales decrease participants' ability to manipulate control levers and thereby test their hypotheses. The nature of our treatment simultaneously facilitated the use of a particular time scale focus and set a minimum control time scale. If the causal relations we propose are correct, changing our treatment
simultaneously facilitated one requirement of learning and constrained another. Because limits on either the ability to reflect or test constrain learning, the combination of intermediate time scales of perspective and control would result in more system understanding than very long or short time scales. This is consistent with our observations (see Figure 5). Apparently, only when perspectives are long enough to allow reflection about delayed system responses and control time scales are short enough to allow testing of suspected causal relations can participants improve their causal understanding.

Control time scales can also influence learning through the allocation of cognitive capacity. Participants using the shortest time scales may be lured by their high level of control to concentrate on maximizing system performance with frequent system adjustments and ignore system understanding, a behavior we refer to as the "video game syndrome," which also has been observed by other researchers (Machuca et al., 1998; Paich & Sterman, 1993). This intense focus on the next adjustment could leave participants with inadequate cognitive capacity to develop the longer perspective needed to improve their causal understanding of the system as a whole (Brehmer, 1998). Therefore, very short control time scales would be expected to be correlated with less learning.

The interaction between perspective length and the control time scale explains our system understanding improvement data better than our original hypothesis. Our improved hypothesis is that the time scale focus length and time scale of control interact in a multiplicative function to influence system understanding improvement. Longer time scale focus lengths positively influence the understanding of the dynamics of systems, whereas longer time scales of control negatively influence the ability to manipulate the system and test hypothesized impacts of policy alternatives. This hypothesis can help explain why participants with the shortest and longest time scale treatments had less causal improvement than those with moderate time scales but increasing time scales consistently degraded performance. Additional research that distinguishes between the effects of perspective length and the time scales of control is needed to test our improved hypothesis.

Conclusion

Understanding how user time scale focus length affects learning can provide the basis for improved MFS design. An experiment was used to measure learning with causal understanding and managerial performance. Results do not statistically support the hypothesis that a long time scale focus increases understanding of structural relationships more than a short time scale focus. However, our analyses suggest that features other than time scale focus influence learning. Causal distance and timing of impact effects were not supported as important factors. An improved hypothesis that the time scale of control and time scale of focus interact multiplicatively to influence the ability of participants to learn experimentally was developed based on further analysis and the results of a subsequent experiment.
We conclude that increasing the time scale of focus does not monotonically increase system understanding improvement as suggested in the literature. Based on our improved hypothesis, we also conclude that to understand their interaction, the relationship must be described through the cognitive mechanisms of the method of learning being used. In the case of experimental learning, the relationship must distinguish between the impacts of perspective length and the flexibility of control on understanding, hypothesis testing, and managerial performance. The nature of these mechanisms requires an understanding of the interaction of these factors as well as understanding them separately.

This research has tested a hypothesis that is based on a fundamental system dynamics principle that perceiving delayed system responses increases system understanding. In doing so, we have improved the design and tools for measuring learning in dynamic systems. Our analysis of causal structure understanding along the causal distance and timing of impacts dimensions expands the dimensions for investigating learning about dynamic systems. We provide evidence that the time scale of control is an important factor affecting learning as well as performance. Although we consider our conclusions to be too preliminary for use as MFS design guidelines, they point to time scales as an important feature to be considered by MFS designers and to the cognitive requirements of specific learning methods as MFS design guidelines. DSS researchers also benefit by the identification and more specific description of several important aspects of DSS design for future research.

Future research can test our fundamental hypothesis with experiments that further isolate the effects of time scale focus length from other factors that influence learning. Our measurements of learning can be expanded to include additional types of structure understanding, behavior patterns, and the interaction of structure and behavior. Our improved hypothesis can be tested by investigating the relationship between time scales of control and learning or the interaction of the time scales of control and focus. Improved understanding of how time scales affect system understanding can provide a basis for improved MFS design and the development of DSSs that can effectively facilitate learning about dynamic systems.

Note

1. POWERSIM is a registered trademark of ModelData of Bergen, Norway.

References


O'Keeffe, R. (1987). What is visual interactive simulation (and is there a methodology for doing it right?). Proceedings of the 1987 Winter Simulation Conference, Piscataway, NJ.


---

David N. Ford is an assistant professor in the Department of Information Science at the University of Bergen, Norway. Professor Ford holds a Ph.D. from the Massachusetts Institute of Technology and researches the uses of simulation to improve project management and product development.

Dalton E. M. McCormack is a doctoral candidate at the Department of Civil Engineering at the University of Bergen, Norway. He also holds a bachelor's degree (Hon.) in electrical engineering from the Fourth Bay College, Sierra Leone, and a master's degree in philosophy from the University of Bergen, Norway.

ADDRESSSES: David N. Ford, Department of Civil Engineering, Texas A&M University, College Station, TX 77843-3136; fax +1 979-845-6554; e-mail davidford@tamu.edu. Dalton E. M. McCormack, Department of Information Science, University of Bergen, 5020 Bergen, Norway; telephone +47 55-58-41-19; fax +47 55-58-41-07; e-mail daltonm@ifi.uib.no.
CONTENTS

Effects of time scale focus on system understanding
in decision support systems. ...................... David N. Ford
Dalton E. M. McCormack 309

Symposium Issue: System Dynamics and Interactive Learning Environments, Part 2
Guest Editors: Pål I. Davidsen, J. Michael Spector, and Konrad Morgan

A dynamic simulation game (UNIGAME) for strategic university management .................. Yaman Barlas
Vedat Gülçü Diker 331

Group model building:
A decision room approach. ...................... Etienne A. A. Rouwette
Jac A. M. Vennix
Cécile M. Thijsen 359

Learning as acquisition and learning as interaction ..................................... Jan H. G. Klabbers 380

Reports & Communications
The RESTAURANT GAME ...................... Dallas Brozik
Alina Zapalska 407

Simulation/Game
FREELANCE FACILITATORS .................. Sivasailam Thiagarajan 417

Simulation/Game Reviews
THRESHOLD COMPETITOR .................. John C. Schafer 422
SERVICESIM .............................. Albert D. Widman 426

Association News & Notes
ABSEL ........................................ John K. Butler, Jr. 429