Testing the Effects of a System Dynamics Decision Aid on Mental Model Accuracy and Performance on Dynamic Decision Making Tasks

Michael Shayne Gary
Australian Graduate School of Management
University of New South Wales
Sydney NSW 2052 Australia
email: sgary@agsm.edu.au
Telephone: +61 2 9931-9247

Robert E. Wood
Australian Graduate School of Management
University of New South Wales
Sydney NSW 2052 Australia
email: rwood@agsm.edu.au
Telephone: +61 2 9931-9180

ABSTRACT

Previous studies have suggested that decision aid support can positively affect performance on dynamic decision making tasks. However, few studies have examined the effects of decision aid support on the cognitive mechanisms underpinning performance. This study tested the relationships between decision aid support (in the form of causal loop diagrams), task complexity, cognitive load, mental model accuracy and performance, using a product lifecycle management simulation. Results indicate that task complexity and decision aid support are significant predictors of mental model accuracy, and that decision aid support moderates the relationship between task complexity and mental model accuracy. In addition, task complexity and mental model accuracy are significant predictors of performance. Our findings regarding the beneficial impact of decision aid support on mental model accuracy and on incremental gains in learning and performance highlight the importance of understanding the underlying cognitive mechanisms at work. Designing more effective decision aids to enhance the development of accurate mental models is one path with a great deal of potential to improve performance in dynamic decision making environments.

Keywords: dynamic decision making, cognitive load, mental models

Much of the work in system dynamics is targeted towards helping managers and policymakers learn about and adopt more effective policies in the dynamic decision making environments in which they operate. Dynamic decision making is part of our everyday lives and is particularly pervasive in organizational contexts. Specific examples include managing a team, launching a new product, managing a multi-stage supply chain, or deploying resources during emergencies. In such information feedback contexts, decisions are often interdependent and the environment changes as a consequence of the decision maker's actions (Brehmer, 1992; Edwards, 1962; Forrester, 1961).

A key assumption in system dynamics is that the policies and decisions managers and policymakers typically adopt when managing complex systems are often responsible for many of the problems experienced in business and social systems. A large volume of work provides support for this position (for just a few examples see, Forrester, 1959; Forrester, 1961; Forrester, 1969; Lyneis, 1980; Meadows, Meadows, Randers, & Behrens, 1971; Meadows, 1970; Sterman, 2000). In addition, the results from a growing body of experimental research also supports this position and indicates that decision making and performance are far from optimal in complex, dynamic decision environments (Langley & Morecroft, 2004; Paich & Sterman, 1993; Sengupta & Abdel-Hamid, 1993). The experimental research also finds that learning plateaus rapidly and experience does not seem to improve the effectiveness of the decisions in such environments after the initial learning plateau (Paich & Sterman, 1993; Langley & Morecroft, 2004). Such experimental work provides crucial evidence in support of widespread decision making errors and misperceptions of feedback in complex systems.

Experimental work on dynamic decision tasks also enables researchers to test the efficacy of interventions aimed at improving mental models, decisions making and performance in complex systems. A variety of interventions are used in the system dynamics community to improve mental models and decision making. Examples include modeling projects with client managers and policymakers, identification of generic structures underpinning common business and social problems, system dynamics courses in K-12 education and universities, qualitative systems thinking interventions focused on eliciting causal feedback structure, and the spread of simulation microworlds into the curriculum of a wide range of university courses. Each of these interventions can be subjected to controlled testing in experiments.

A number of previous studies have suggested two related mechanisms that may underpin poor performance in complex tasks: 1) cognitive load and 2) mental model accuracy. Several studies have used these constructs as post hoc explanations of poor performance, but have not subjected them to empirical testing (Dielh & Sterman, 1995; Paich & Sterman, 1993). The rationale is that humans experience high levels of cognitive load when solving complex problems, and that high cognitive load impairs the formation of accurate mental models of the decision environment (Sweller, 1988; Paich & Sterman, 1993). Deficient mental models are, following this logic through, responsible for decision makers' poor decisions and performance in complex decision environments.

This study builds on previous experimental work in dynamic decision making to test the impact of providing decision makers with information about the feedback structure of a complex system in which they are operating. Specifically, we examine the relationships between decision aid support (providing information about the feedback structure), task complexity, cognitive load, mental model accuracy and performance, using a modified version of Paich and Sterman's (1993) B&B product lifecycle management simulation. We investigate the underlying cognitive mechanisms of poor decision making and performance. In the next section, we review the relevant theory and outline the hypotheses tested in this study. Subsequently, we describe our methods and then present the empirical results. We conclude by discussing how our findings contribute to prior dynamic decision making research and identify areas requiring future study.

THEORY AND HYPOTHESES

Mental Models

Mental models are simplified representations of reality that are encoded within the mind of an individual. Doyle and Ford (1998) proposed the following definition of mental models:

> "Mental models are relatively enduring and accessible, but limited, internal conceptual representation of a system" (Doyle & Ford, 1998)

Causal attributions are a central feature of mental models (Sterman, 1994). Some definitions of mental models also include the concept of a set of stored procedures in the form of heuristics or automatic-task responses, which are closely related to the policies or decision rules incorporated into system dynamics models. This broader definition suggests that mental models are composed of both declarative and procedural knowledge (Sweller, 1988). Mental models that are composed of high-quality information and that include rich and sophisticated linkages, help decision makers attend to the most important cues in their environment, facilitate effective encoding and retrieval of information, and guide effective problem solving processes (Eden & Spender, 1998; Hodgkinson, 2003; Hodgkinson & Sparrow, 2002; Walsh, 1995).

A large number of previous studies have asserted that people develop incomplete and inaccurate mental models of dynamic decision environments resulting in misperceptions of feedback between decisions and the environment (Diehl & Sterman, 1995; Paich & Sterman, 1993; Sengupta & Abdel-Hamid, 1993; Sterman, 1989). Prior research has identified two specific deficiencies in mental models due to misperceptions of feedback: 1) misperceptions of feedback structure and 2) misperceptions of feedback dynamics (Diehl & Sterman, 1995; Paich & Sterman, 1993; Sterman, 1989). Misperceptions of feedback structure arises because the underlying causal relationships of complex decision environments–which often include time delays, feedback effects, stock accumulation processes and nonlinearities–are difficult to *map* and *integrate* into decision makers mental models (Paich & Sterman, 1993; Sterman, 1989). Misperceptions of feedback dynamics arises because decision makers are generally not capable of accurately inferring the behavior of dynamic systems to determine the consequences of their decisions (Diehl & Sterman, 1995; Sterman, 1989).

Cognitive Load

The cognitive load generated by a task refers to the total amount of mental activity it imposes on working memory at a given point in time. Drawing upon assumptions made on human cognitive architecture, Sweller, van Merrienboer, & Paas (1998) stipulate that there are three different types of working memory load. The first is *intrinsic cognitive load* – which are demands that are imposed by the inherent complexity of the materials being learned. The second, *extraneous cognitive load, is* load that is induced by the format in which the information is presented. Extraneous load is regarded as unnecessary for learning as it simply depletes working memory capacity and does not promote schema acquisition or automation. Finally, cognitive load that enhances learning by promoting schema formation and automation is called *germane cognitive load*.

Information may only be stored in long term memory after first being generated and attended to in working memory (Atkinson & Shiffrin, 1968; Sweller, 1988). Working memory however, is extremely limited in capacity and when the information to be learned is complex in nature (i.e. includes a large number of interacting elements), the cognitive load imposed can quickly exceed working memory capacity (i.e. cognitive overload) disrupting the formation of accurate mental models (Sweller, 1988).

Cognitive Feedback: Decision Aid Support

Only a few studies have investigated the effects of including cognitive feedback in the form of decision aids in complex dynamic environments. There are various forms of feedback, but the three pertinent to dynamic decision making are a) outcome feedback, b) feedforward, and c) cognitive feedback. Outcome feedback is information one receives regarding their performance after each decision trial (e.g. net profit earned). Feedforward on the other hand is information intended to reduce the cognitive demands of a decision maker by providing them with a basic set of decision heuristics (e.g. clearly defined computational steps) to follow when making subsequent decisions within the task environment. Finally, cognitive feedback is information intended to improve decision making by enhancing a person's comprehension and knowledge of the task structure (Blazer et al., 1989; Hammond et al., 1975). The are various components of cognitive feedback, but one of the most effective relates to task information, or information about the relationships between important target variables in the decision environment (Blazer et al. 1989; Sengupta & Abdel-Hamid, 1993).

In a study conducted by Sengupta and Abdel-Hamid (1993), the three forms of decision feedback aids were compared using a software project management simulation to test which would be the most helpful in improving performance. The results from their study revealed that participants supplemented with cognitive task information feedback depicting the relationships among various cues of the task, performed best relative to those receiving only feedforward or outcome feedback.

In another study, Langley and Morecroft (2004) also examined ways to improve performance through online information feedback. Their results indicate that subjects provided with cognitive feedback in the form of a causal map depicting the relationships between the key decision variables, performed significantly better throughout the task than the groups provided only with outcome feedback. In addition, their study showed that even when cognitive feedback is removed in later trials performance increments are sustained, suggesting that participants form an internal representation of the information in the task structure despite limited exposure. It is interesting to note however, that although the group that received the cognitive feedback performed significantly better than the control group in the initial trials, in the latter trials performance was almost equivalent. Despite the similar performance at the end, the results still suggest that cognitive feedback may have important beneficial effects on learning and performance (Langley and Morecroft, 2004).

Researchers argue that cognitive feedback is more effective than feedforward information or outcome feedback in isolation, because it allows decision makers to form appropriate mental models of the system and enhance their ability to identify trends and detect any changes underlying the system (Sengupta & Abdel-Hamid, 1993). In turn this allows them to update their strategies and improve performance. In contrast, outcome feedback presented trial-by-trial may not facilitate the formation of accurate mental models since decision makers must draw accurate causal inferences in a complex environment. Also, feedforward information or heuristics may not be as effective as cognitive feedback since feedforward information can be rendered ineffective by changes in the decision environment (Brehmer, 1990; Sengupta & Abdel-Hamid, 1993).

Hypotheses

This study compares the effects of decision aid support during the learning phases of a complex task on cognitive load, mental model formation and dynamic decision making performance. We focus specifically on cognitive feedback in the form of a causal loop diagram because the results from Langley and Morecroft's (2004) study indicate that feedback in this form improves performance more significantly than other feedback formats. Prior research has not examined the effects of decision aids on cognitive load or mental model accuracy to identify the underlying mechanisms for performance differences between groups who received decision aid support versus control groups.

Complex tasks have been shown to induce cognitive load, and therefore we expect cognitive load will be positively associated with higher levels of task complexity. Based on previous dynamic decision making research findings that decision aid support in the form of task information increases performance, we also expect that exposure to a decision aid will reduce the experienced cognitive load. Further, we predict that exposure to a decision aid will moderate the effects of task complexity on cognitive load.

Hypothesis 1: Exposure to a decision aid will reduce participants' cognitive load during a dynamic decision making task. Also, cognitive load will be positively associated with higher levels of task complexity, and exposure to a decision aid will moderate the effects of task complexity on cognitive load.

Research based on cognitive load theory has shown that a high level of cognitive load tends to be detrimental to task performance (Sweller, 1988). We predict that mental model accuracy will be negatively related to higher levels cognitive load. Specifically, decision makers who experience higher levels of cognitive load will form less accurate mental models than those who experience lower levels. Prior research also indicates mental model accuracy is negatively related to task complexity (Gary & Wood, 2006), and we also predict that mental model accuracy is negatively related to the level of task complexity. Previous research has also found that providing task structure information as a decision aid results in higher task performance, and there has been speculation that a decision aid providing task structure information will increase mental model accuracy (Sengupta & Abdel-Hamid, 1993). Therefore, we predict that exposure to a decision aid will increase mental model accuracy and that these relationships will be significant after controlling for self efficacy[1]. We also predict that decision aid support provided in the form of a causal loop diagram will moderate the effects of task complexity on mental model accuracy.

Hypothesis 2: Mental model accuracy will be negatively related to cognitive load and task complexity, and positively related to exposure to a decision aid providing task structure information after controlling for self efficacy.

Hypothesis 3: Decision aid support will moderate the effects of task complexity on mental model accuracy.

Past studies have shown that decision aid support and task complexity both impact decision making performance. However, prior research has not examined the interaction between these variables. Based on cognitive load theory arguments that decision aid support can alleviate the cognitive demands induced by high complexity task environments, we predict that decision

[1] General cognitive ability was assessed using participants' university admissions test as a broad indicator of an individual's capacity to learn. The inclusion of the university admissions test would have enabled examination of the relationships between cognitive ability, cognitive load, mental model accuracy and performance. Due to extensive missing data this variable was subsequently excluded from the analyses. The missing data was mainly due to the fact that many participants were from other states and countries that did not have an equivalent test.

aid support in the form of task structure information will moderate the effects of task complexity on performance.

> *Hypothesis 4: Decision aid support will moderate the effects of task complexity on performance.*

Recent research indicates that mental model accuracy is positively related to performance on complex tasks (Gary and Wood, 2006). We predict that transfer performance will be positively related to mental model accuracy and decision aid support, and negatively related to task complexity and cognitive load after controlling for self-efficacy.

> *Hypothesis 5: Performance will be positively related to mental model accuracy and decision aid support, and negatively related to cognitive load and task complexity after controlling for self efficacy.*

The next section outlines our experimental design, describes the task and measures, and summarizes the data collection procedures.

METHOD

Experimental Design

The study employed a 2x2x(6) experimental design with repeated measures. The two between group factors were: task complexity (low/high) and decision aid support (yes/no). Participants were randomly allocated to the four groups. The dependent variables consisted of: cognitive load, mental model accuracy, and performance on the simulation task. All participants completed a learning phase of three trial blocks consisting of 120 decision trials and subsequently completed an immediate transfer phase of another three trial blocks of 120 decision trials.

Participants

Second and third year undergraduate students enrolled in a Bachelor of Commerce program (including Accounting, Finance, Management, and Marketing) with no prior experience on the simulation were invited to participate. The 99 participants were randomly assigned to one of the four experimental conditions (Low Complexity with Decision Aid: 28, Low Complexity Control: 23, High Complexity with Decision Aid: 26, High Complexity Control: 22). The sample consisted of 40 men and 59 women, with an average age of 21 years. Each participant received three movie tickets at the end of the study for their full participation.

Task Environment

Simulation. The study was conducted using an interactive computer-based management simulation. There were two different versions of the simulation–high and low complexity–and for each level of complexity there was a decision aid and no decision aid condition. During the simulation, participants assumed the role of the managing director of a company and their task was to manage the launch of a new product through a forty quarter (10 year) lifecycle. This simulation has been used extensively in previous research (Gary & Wood, 2006; Paich & Sterman, 1993). Participants in the low complexity condition managed a monopoly business with two decision variables: price and target capacity. Those in the high complexity version managed three decision variables: price, target capacity and also marketing expenditure (see Figure 1). The high complexity version also had a competitor sector, making it more difficult for the participant to earn profits in the market. Participants made their decisions for each of the variables every quarter and entered them directly before deciding when to advance to the next quarter (by pressing the Simulate button).

Insert Figure 1 about here

The goal of the management simulation was to maximize cumulative profit from the sales of the product through a forty quarter lifecycle. After each decision trial, outcome feedback was provided in both tabulated and graphical format to control for the differential effects that feedback format may have on decision makers (Atkins, Wood, & Rutgers, 2002).

Interactive Decision Aids. Two separate decision aids corresponding to the two complexity levels of the study were developed. Each decision aid consisted of an overall causal map (or causal loop) diagram consisting of words and arrows depicting the direction of influence between the full set of variables in the simulation. The words and arrows were color coded to help participants interpret the task structure information. In addition, each variable and causal arrow was hyperlinked so that a mouse click brought up a pop-up window with a brief description of the variable or of the relationship implied by the arrow and a short explanation about how two variables are related (see Figure 2). The decision aid was embedded within the simulation and participants in the experimental conditions could access it at anytime during the initial learning phase of the experiment.

Insert Figure 2 here

Measures

Cognitive Load Ratings and Decision Aid Referral. Cognitive Load was assessed using a validated nine-point subjective rating scale adopted from Kalyuga, Chandler, Tuovinen & Sweller (2001). Participants were asked to estimate how easy or difficult they found the simulation task by circling one of the answers ranging from 1: Extremely Easy to 9: Extremely Difficult. The higher the score, the higher the estimated mental load. Participants were also asked a few questions with regards to the amount information they were required to incorporate and process, as well as how stressful they found the task overall.

Using the same scale format as the cognitive load measure, the decision aid referral measure asked participants to indicate how easy or difficult it was to process the information in the decision aid. In addition, the questionnaire also asked participants to indicate how often they referred to the decision aid during the simulation and how helpful they found it.

Mental Model Accuracy. After the learning phase, participants' mental models of the task were assessed using a knowledge test. One set of questions tested participants' recall of bivariate causal relationships between pairs of variables from the management simulation. A second set of questions tested participants' ability to infer the dynamics of small sets of interdependent variables from the new product launch simulator. The knowledge test was designed to assess some aspects of the two components of misperceptions of feedback discussed previously. Accuracy of an individual's mental model of the decision environment is therefore a function of: (1) the perceived causal relationships between pairs of variables in the decision environment inferred through their experience in the task domain, and (2) their ability to infer the dynamics of small sets of interdependent variables in the decision environment.

In the first set of questions, participants' knowledge and recall of the bivariate causal relationships between pairs of variables from the management simulation were tested, including the sign or polarity if there was a relationship. The questions covered the exhaustive set of actual relationships in each of the complexity conditions along with several items for which no relationship existed in the decision environment. Participants in both complexity conditions answered 30 items on the relationships between variables that were common to both decision environments. Participants in the high complexity condition answered a further 24 items relating to the additional variables and relationships in the high complexity condition. Appendix A provides a segment of the instructions along with the first three items of this set of questions.

In the second set of questions, participants' knowledge of the relationships between a small set of variables in the management simulation and also their ability to infer the dynamics of this set of variables was tested. Each question presented the graph of one or two variables over time from the task, and asked subjects to choose from a multiple choice of answers for the evolution of another variable in the task. Participants must draw on their experience with the management simulation and their knowledge of the relationships between variables in order to determine how the dynamic behavior of the first variable or variables impacts the dynamic behavior of another variable. Appendix B provides a segment of the instructions along with one example from this set of questions.

Each item on the knowledge test was scored as correct or incorrect for each participant. There are nine possible ways to answer each influence diagram question given that each item could be answered using one directed arrow, two directed arrows in a closed feedback loop, or by writing NONE if there was no direct causal relationship between the two variables. Therefore, a random answer strategy on the questionnaire would result in a score of 11% accuracy. Each graphical scenario question had four multiple choice answer options, and therefore a random answer strategy would result in a score of 25% accuracy. Mental Model Accuracy was the sum of the percentage of items on the knowledge test answered correctly for each of the two sets of knowledge questions. That is the total number of correct items divided by the total number of questions for each of the two types of questions on the test-the bivariate causal relationships and dynamics in graphs over time. The possible scores range from 0 to 2, where a score of 2 indicates perfect knowledge of the causal structure and dynamic behavior of small sets of variables in the decision environment.[2]

[2] Another questionnaire was administered to participants after the learning phase that tested their knowledge of bivariate relationships in the management simulation using a True/False format. Items from the True/False questionnaire were used to cross-validate the reliability of responses obtained on the first set of questions using the Influence Diagram answer format.

Performance

The cumulative profit at the end of each forty-quarter trial block was used as the performance indicator. Because the potential achievable cumulative profit was different in the high and low complexity conditions, subjects' raw scores were assessed relative to high performance benchmarks calculated using a standard computational technique for both conditions. The cumulative profit benchmarks were found through single point optimization using a modified Powell search implemented in Vensim simulation software. The benchmarks were not global optima for the task, but instead were simply a consistently calculated high performance benchmark.

Control Variable: Self Efficacy

Perceived self efficacy was measured using a 10-item scale covering a broad range of activities participants needed to manage throughout the simulation. The format followed the approach presented by Bandura (1997), which has been validated in numerous empirical studies. For each item, participants first indicated whether or not they understood what was required to manage a specific activity (yes or no), and then to record their confidence in their capabilities on a 10-point rating scale (ranging from 1: Very low confidence to 10: Very high confidence). Perceived self efficacy was included to ensure that differences in the performance and mental models of the participants' were not solely attributable to motivational differences.

Procedure

Participants were tested in small groups with each individual allocated to an individual computer. The test session lasted between two and a half to three hours depending on the pace of the participant. Prior to sitting the simulation, participants in the decision aid conditions were required to work through a series of exercises to get familiar with the decision aid. The purpose of the exercises was to help train the participants on how to effectively use the decision aid. After completing the decision aid exercises, participants then proceeded to the learning phase of the experiment. The learning phase consisted of *three blocks* of 40 decision trials each, 120 decision trials in total, to enable participants to learn about and become familiar with the task environment. After completing the learning phase, participants were asked to complete a series of questionnaires assessing their cognitive load, self efficacy, and mental models of the task. Those in the decision aid conditions received an additional decision aid reference questionnaire. Following the questionnaires, participants proceeded to the immediate transfer testing phase and completed three more blocks of 40 decision trials each. The decision aid was removed during this final transfer phase in order to examine the effects of its inclusion in the initial learning trials on performance in the subsequent trials.

The procedure for the two control conditions without decision aid support was identical to the decision aid conditions with the only exception being that they did not receive the pre-learning phase decision aid exercises or have access to the decision aid during the learning phase. Participants were under no strict time pressure and completed each phase at their own pace.

Data Analyses

To test Hypotheses 1 and 3, the relationship between decision aid support and task complexity on cognitive load and mental model accuracy was examined using two separate Two-way ANOVAs. The interaction effects allowed us to explore whether decision aid support moderated the effects of task complexity on cognitive load and mental model accuracy respectively.

To test Hypothesis 2, multiple regression analyses was used to examine the unique relationship between cognitive load and mental model accuracy with the effects of complexity, decision aid support, cognitive ability and motivation partialled out.

To test Hypothesis 4, a Two-way MANOVA with repeated measures, was used to assess whether the decision aids moderated the effects of complexity on performance across the six trials.

To test Hypothesis 5, General Linear Models (two fixed factors: complexity and decision aid) with repeated measures on the immediate-transfer phase (trial blocks: 4, 5, and 6) and covariates (cognitive load, cognitive ability, motivation and mental model accuracy) was used to examine the relationship between mental model accuracy and immediate transfer performance.

RESULTS

The means, standard deviations, and correlations for the variables in the study are presented in Table 1. Task complexity was coded so that 0 = low complexity and 1 = high complexity. Decision aid was coded so that 0 = no (i.e. control groups) and 1 = yes.

Insert Table 1 here

Tests of Hypotheses

The effects of decision aid support and task complexity on cognitive load predicted in Hypothesis 1 were partially supported. As shown in Table 2, the main effect for complexity was significant [F (1, 95) = 12.93, $p < .01$, partial $\eta^2 = .12$]. The high complexity groups reported higher levels of cognitive load than the low complexity groups. This is consistent with previous research that shows that tasks which are more complex and difficult in nature tend to induce high levels of cognitive load. This also serves as a manipulation check for the two levels of task complexity. The main effect for decision aid support [F (1, 95) = 2.68, $p = .10$, partial $\eta^2 = .02$] was not significant, but was approaching significance. The interaction between decision aid and task complexity was not significant [F (1, 95) = .04, $p = .83$, partial $\eta^2 = .00$]. On average across all groups, participants rated the difficulty level of the decision aid to be 4.78 out of 9.0. This indicates the decision aid itself induced cognitive load and may explain why providing decision aid support did not impact perceived cognitive load ratings of the task. Therefore, although the results do not support our hypothesis that decision aid support will moderate the effects of task complexity on cognitive load, a significant main effect of complexity on cognitive load did emerge.

Insert Table 2 here

The relationships between decision aid support, task complexity, cognitive load and mental model accuracy proposed in Hypothesis 2 were all supported apart from one relationship discussed below. Table 3 provides the unstandardised and standardized regression coefficients for the model. Task complexity ($t = 2.00$, $p < .05$), decision aid support ($t = 2.39$, $p < .05$), and self efficacy ($t = 4.68$, $p < .001$) were found to contribute significantly to mental model accuracy [$F(4, 90) = 8.35$, $p < .001$]. The model accounted for about 24% of the variation in mental model accuracy (*Adjusted* $R^2 = .24$). Cognitive load ($t = 1.23$, ns) was not a significant predictor of mental model accuracy. This is the first study to explicitly test whether providing decision aid support in the form of task structure information leads to the formation of more accurate mental models in complex dynamic decision environments. Our results indicate providing decision aid support in the form of a causal loop diagram does indeed lead to more accurate mental models.

Insert Table 3 here

The relationships between decision aid support, task complexity and mental model accuracy proposed in Hypothesis 3 were supported. As shown in Table 4, the interaction between decision aid support and task complexity was significant [$F(1, 94) = 3.92$, $p = .05$, partial $\eta^2 = .04$]. Participants in the low complexity version of the task with decision aid support developed more accurate mental models than participants in the high complexity decision aid support condition. Overall, decision aid support does moderate the effects of task complexity on mental model accuracy.

Insert Table 4 here

The relationships between decision aid support, task complexity and performance proposed in Hypothesis 4 were not supported. Figure 3 shows the mean performance across the six trial blocks for the four groups. Trial Blocks 1-3 were the learning phase and trial blocks 4-6 were the immediate transfer phase. As can be seen in Table 5, the interaction effect [$F(1, 88) = .00$, $p = .97$, partial $\eta^2 = .01$] between decision aid support and task complexity on performance was not significant. Therefore, there is no evidence to suggest that decision aid support moderates the effects of task complexity on performance.

Insert Figure 3 here

Insert Table 5 here

The relationships between mental model accuracy, decision aid support, cognitive load, task complexity and performance proposed in Hypothesis 5 were partially supported. The results presented in Table 6, indicate that mental model accuracy and task complexity [F (1, 85) = 39.98, $p < .001$] are significant predictors of performance in the immediate-transfer phase of the experiment [F (1, 85) = 8.70, $p < .01$], after controlling for self efficacy. However, decision aid support [F (1, 85) = 0.05, *ns*] and cognitive load [F (1, 85) = 0.64, *ns*] were not significant predictors of performance. Therefore, mental model accuracy is positively related to subsequent performance and participants in the low complexity conditions performed significantly better than those in the high complexity conditions in the immediate-transfer phase of the study.

Insert Table 6 here

DISCUSSION

The aim of the current study was to investigate the mechanisms responsible for poor performance relative to the potential achievable levels on complex, dynamic decision-making tasks. In particular, we investigated the relationships between decision aid support, cognitive load, mental model accuracy, task complexity, and performance. Overall, the results indicate that task complexity has a significant impact on cognitive load, mental model accuracy, and performance. Decision aid support is also a significant predictor of mental model accuracy; those with exposure to the decision aid develop more accurate mental models than those in the control groups. Mental model accuracy also has a significant effect on performance. Participants with more accurate mental models after the learning phase of the study outperformed those with less accurate mental models in the transfer phase.

Past studies have suggested that high task complexity can have negative effects on cognitive load (e.g. Sweller, 1988, Merrienboer & Sweller, 2005). In line with the findings from past studies, the present study did find that task complexity itself is a strong predictor of cognitive load. The results emphasize the importance of finding alternative decision aid support or feedback interventions that reduce cognitive load in dynamic decision making tasks.

Our results also indicate that task complexity is a significant predictor of mental model accuracy. These findings are consistent with cognitive load theory (Sweller, 1988) and also with the misperceptions of feedback arguments about the detrimental effects of complex decision environments on the formation of accurate mental models (Dielh & Sterman, 1995; Paich & Sterman, 1993). Further, our results indicate that participants who received decision aid support, in the form of task structure information in a causal loop diagram, formed more accurate mental models than those who did not. This new finding is important because it demonstrates supplemental decision aid support that illuminates the key relationships between variables in the task environment can lead to the formation of more accurate mental models in dynamic decision making tasks.

Regarding the relationship between cognitive load and mental model accuracy, there was no evidence in this current study that cognitive load is a predictor of mental model accuracy. These results contradict past findings that experiencing high levels of cognitive load may impair the formation of accurate mental models (Sweller, 1988, Merrienboer & Sweller, 2005). However, minimal empirical research has attempted to explicitly examine the relationship between cognitive load and mental model accuracy. There is evidence in our study that the decision aid itself induced a high degree of cognitive load which may explain these results. Far simpler decision aids may be required for initial learning with a progressive sequence of more complex decision aids provided as decision makers learn about the problem domain. The counterintuitive results of the current study may have been due to the participants experiencing different forms of cognitive load as stipulated by Sweller et al. (1998). For example, those who found the decision aid support useful may have experienced high levels of germane cognitive load and formed more accurate mental models. On the other hand those who did not find it useful may have experienced high levels of extraneous cognitive load and formed less accurate mental models. The differential effects of the different types of cognitive load on mental model accuracy formation would have made it difficult to detect any significant effects. Further research should try to delineate and measure the different forms of cognitive load.

Our results indicate that performance is negatively related to task complexity and this finding is consistent with previous studies in which high complexity tasks have detrimental effects on performance (Diehl & Sterman, 1995; Paich and Sterman, 1993). Diehl and Sterman (1995) suggested that future studies should examine the use of decision aids to help individuals overcome the poor performance outcomes that are often associated with complex tasks, but there have only been a handful of studies attempting to build on our understanding in this area (Sengupta & Abdel-Hamid, 1994; Langley and Morecroft, 2004; Gonzalez, 2005). This study tested the impact of providing decision aid support on mental model accuracy and performance. Providing decision aid support had a significant impact on mental model accuracy. We did not find a significant relationship between decision aid support and performance after controlling for other variables. However, mental model accuracy may mediate the relationship between decision aid support and performance.

As in previous research, our current study also found that average performance improves initially, but then plateaus quite rapidly at a level far below the potential achievable level. These findings are consistent with previous research findings that learning in dynamic decision making tasks tends to plateau quickly (Paich & Sterman, 1993; Langley & Morecroft, 2004). To account for these performance plateaus, prior studies have suggested that participants may devise simple decision rules based on a small number of cues in the decision environment, and thereafter apply these decision rules somewhat automatically (Dielh & Sterman, 1995; Paich & Sterman, 1993; Sterman, 1989). This would be consistent with a large body of research focusing on the acquisition of skills, which suggests novices and generally anyone solving a novel complex problem tend to automate decision and action rules prematurely resulting in low levels of performance compared with potentially achievable levels (Anderson, 1982, 1987; Ericsson, 1993, 2003). Future research could collect detailed process measures of decision making trial by trial to shed light on this issue and test the propositions about decision rule automaticity.

Given the positive relationship between mental model accuracy and performance, our findings regarding the beneficial impact of decision aid support on mental model accuracy and on incremental gains in learning and performance are particularly important. Prior experimental research findings provide evidence of widespread decision making errors and misperceptions of feedback in complex systems (Diehl & Sterman, 1995; Paich & Sterman, 1993; Sterman, 1989). Experimental work on dynamic decision making enables researchers to test the efficacy of interventions aimed at improving mental models, decision making and performance in complex systems. This study tested the efficacy of just one system dynamics intervention under controlled conditions. Providing decision makers with causal loop diagrams of complex decision environments does improve mental model accuracy. However, there is still massive scope for performance improvements; especially at high levels of complexity. This combination of results provides strong motivation for further research in finding more effective decision aid designs to supplement dynamic decision making. Future research needs to identify better interventions to enhance the formation of accurate mental models by reducing cognitive load and preventing premature automation of decision rules.

REFERENCES

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369-406.

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem situations. *Psychological Review*, 94(2), 192-210.

Atkinson, R.C. & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds). *The Psychology of Learning and Motivation (Vol 2)*. London: Academic Press

Atkins, P. W. B., Wood R. E., & Rutgers, P. J. (2002). The effects of feedback format on dynamic decision making. *Organizational Behavior and Human Decision Processes, 88,* 587-604.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Blazer, W. K., Doherty, M. E., & O'Connor, R. O. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin, 106,* 410-433.

Brehmer, B. (1992). Dynamic decision making: human control of complex systems. *Acta Psychologica, 81,* 211-241.

Diehl, E., & Sterman, J. (1995). Effects of feedback complexity on dynamic decision making. *Organizational Behavior and Human Decision Processes, 62, 198-215.*

Doyle, J. K., & Ford, D. N. (1998). Mental Models Concepts for System Dynamics Research. System Dynamics Review, 14(1), 3-29.

Eden, C., & Spender, J.-C. (Eds.). (1998). *Managerial and organizational cognition: Theory, methods, and research*. London: Sage.

Edwards, W. (1962). Dynamic decision theory and probabilistic information processing. *Human Factors, 4,* 59-73.

Ericsson, A. (2003). *The Acquisition of Expert Performance as Problem Solving*. In J. E. Davidson & R. J. Sternberg (Eds.), The Psychology of Problem Solving. Cambridge:

Cambridge University Press.

Ericsson, K. A., Krampe, T. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100, 363-406.

Forrester, J. W. (1959). Advertising: A Problem in Industrial Dynamics. Harvard Business Review, 37(2), 100-110.

Forrester, J. W. (1961). Industrial Dynamics. Cambridge MA: Productivity Press.

Forrester, J. W. (1969). Urban Dynamics. Cambridge MA: Productivity Press.

Gary, M. S., & Wood, R. E. (2006). *Mental models, decision making and performance in complex dynamic environments*. Manuscript submitted for review.

Gonzalez, C. (2005). Decision support for real-time, dynamic decision-making tasks. *Organizational Behavior and Human Decision Processes*, 96, 142-154.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis (sixth edition).* New Jersey: Pearson.

Hammond, K., Stewart, T., Brehmer, B., & Steinmann, D. (1975). Social judgment theory. In M. Kaplan and S. Schwartz (Eds). *Human Judgment and Decision Processes.* New York: Academic Press.

Hodgkinson, G. P. (2003). The interface of cognitive and industrial, work and organizational psychology. *Journal of Occupational and Organizational Psychology*, 76, 1.

Hodgkinson, G. P., & Sparrow, P. (2002). *The competent organization: a psychological analysis of the strategic management process*. Buckingham and Philadelphia: Open University Press.

Kalyuga S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13,* 351-371.

Langley, P. A., & Morecroft, J. D. W. (2004). Performance and learning in a simulation of oil industry dynamics. *European Journal of Operational Research, 155,* 175-732.

Lyneis, J. M. (1980). Corporate Planning and Policy Design. Portland, OR: Productivity Press.

Meadows, D. H., Meadows, D. L., Randers, J., & Behrens, W. W., III. (1971). The Limits to Growth. London: Pan Books, Ltd.

Meadows, D. L. (1970). Dynamics of Commodity Production Cycles. Cambridge MA: Productivity Press.

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

Sengupta K, & Abdel-Hamid, T. K. (1993). Alternative conceptions of feedback in dynamic decision making environments: An experimental investigation. *Management Science, 39,* 411-428.

Sterman, J. D. (1989). Misperceptions of feedback in dynamic decision making. *Organizational Behaviour and Human Decision Processes, 43,* 301-335.

Sterman, J. D. (1994). Learning in and about Complex Systems. System Dynamics Review, 10(2/3), 291-330.

Sterman, J. D. (2000). Business Dynamics: Systems Thinking and Modeling for a Complex World: McGraw-Hill.

Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science, 12,* 257-285.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology, 10,* 251-295.

van Merrienboer, J. J. G., & Sweller J. (2005), Cognitive load theory and complex learning: recent developments and future directions. *Educational Journal Review, 17,* 147-177.

Walsh, J. P. (1995). Managerial and organizational cognition: Notes from a trip down memory lane. *Organization Science*, 6, 280-321.

Appendix A

An excerpt from the set of Influence Diagram Questions on the Knowledge Test

$X \xrightarrow{S} Y$

This arrow indicates that an increase in X results in an increase in Y above what it would have been (all else equal). On the other hand, a decrease in X results in a decrease in Y below what it would have been (all else equal). X and Y move in the SAME direction.

$X \xrightarrow{O} Y$

In contrast, this arrow indicates X and Y move in the OPPOSITE direction. For example, an increase in X results in a decrease in Y below what it would have been (all else equal). On the other hand, a decrease in X results in an increase in Y above what it would have been (all else equal).

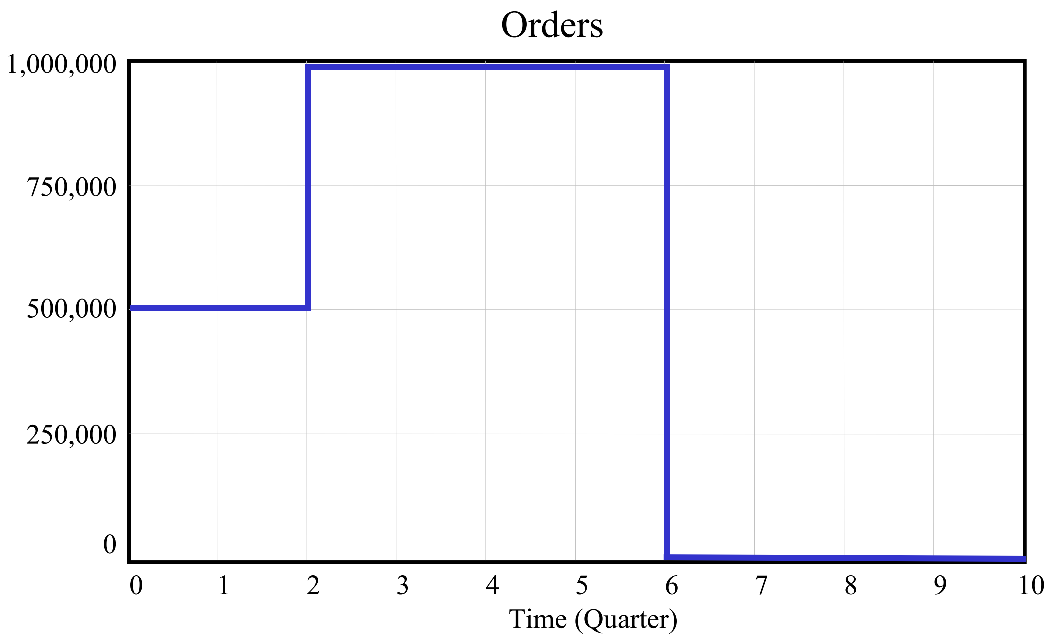
Think about the relationships between these variables that you believe are embedded in the simulator. Relying only on your experience with the simulated firm, draw the appropriate influence arrow(s) for each variable pair and indicate whether the causal influence is in the same or opposite direction using an ‘S’ or ‘O’ at the end of the arrow. Identify any cases in which there is two-way dependency between the variables by drawing the appropriate arrows representing the two-way loop of influence. Focus only on direct relationships and ignore any intervening variables that may result in indirect influence arrows. If there is no direct relationship between the variable pair, write ‘NONE’ between the two variables. If you do not have any idea about the correct answer, then write ‘Do Not Know’ instead of guessing randomly.

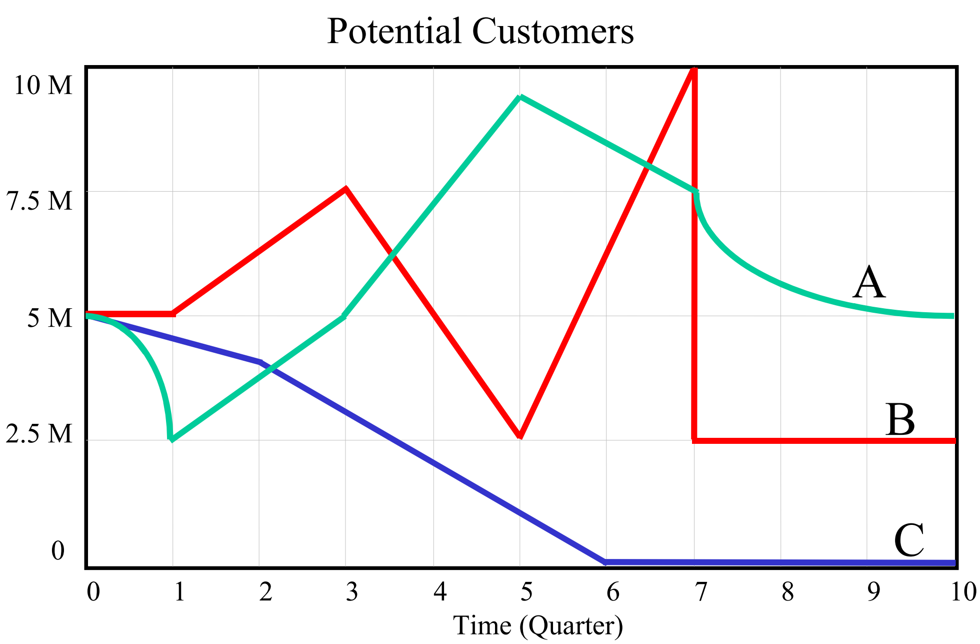
1.	Orders	Backlog
2.	Shipments	Backlog
3.	Backlog	Delivery Delay

Appendix B

Example Graphical Scenario Knowledge Question

Using the time path of Total Industry Orders provided in the top graph below, select the letter of the appropriate time path for Industry Potential Customers on the bottom graph. Circle D if none of the lines in the bottom graph show the correct time path. Assume the initial value of industry Potential Customers is 5 million at Time 0. Also assume that no other variables affect industry Potential Customers over this time horizon.





Answer: A) B) C) D) None of the Above

Table 1. Correlations, means, and standard deviations for the variables in the study

	1	2	3	4	5	6	7	8	9	10	11	12
1. Complexity	1											
2. Decision Aid	.007	1										
3. Performance TB1	-.621**	-.09	1									
4. Performance TB2	-.607**	.015	.675**	1								
5. Performance TB3	-.531**	.043	.568**	.773**	1							
6. Performance TB4	-.548**	.035	.603**	.771**	.785**	1						
7. Performance TB5	-.507**	.214*	.501**	.685**	.797**	.841**	1					
8. Performance TB6	-.493**	.127	.461**	.716**	.847**	.888**	.915**	1				
9. Cognitive Load	.343**	-.158	-.295**	-.225*	-.121	-.153	-.191	-.161	1			
10. Self Efficacy	-.282**	.185	.241*	.267**	.300**	.369**	.365**	.391**	-.359**	1		
11. Mental Model Accuracy	.106	.285**	-.024	.078	.198	.262*	.294**	.301**	-.010	.405**	1	
12. Decision Aid Ref	.046		.118	.034	-.087	.041	-.118	-.044	.199	-.244	-.114	1
Total Sample												
Mean			.07	.28	.35	.38	.41	.46	5.51	5.27	1.06	1.93
Std. Deviation			.57	.42	.40	.38	.38	.39	1.59	1.55	.27	.75
N	99	54	96	98	97	96	98	96	99	96	98	54
Low Complex D.A.												
Mean			.44	.51	.54	.57	.65	.66	4.79	6.10	1.15	1.89
Std. Deviation			.35	.39	.40	.40	.41	.41	1.28	1.23	.26	.11
n	28	28	27	28	27	28	28	28	28	27	28	28
Low Complex												
Mean			.37	.54	.56	.58	.53	.63	5.22	5.20	.89	
Std. Deviation			.45	.40	.38	.36	.35	.38	1.67	1.42	.25	
n	23		23	23	23	23	23	23	23	23	23	
High Complex D.A.												
Mean			-.44	.03	.17	.18	.30	.34	5.81	4.91	1.11	1.96
Std. Deviation			.58	.32	.34	.24	.28	.30	1.35	1.76	.28	.18
n	26	26	22	25	25	24	25	24	26	24	25	26
High Complex												
Mean			-.13	.00	.09	.14	.11	.17	6.36	4.72	1.06	
Std. Deviation			.31	.16	.13	.18	.16	.19	1.67	1.45	.21	
n	22		22	22	22	21	22	21	22	22	22	

** p < .01 (2-tailed).
 * p < .05 (2-tailed).

Table 2. Two-way ANOVA results for the effects of decision aid support and task complexity on cognitive load

Dependent Variable: Cognitive Load

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Corrected Model	34.991(a)	3	11.664	5.23	.002	.142
Intercept	3014.589	1	3014.589	1352.43	.000	.934
Complexity	28.822	1	28.822	12.93	.001	.120
Decision Aid	5.980	1	5.980	2.68	.10	.027
Complexity * Decision Aid	.095	1	.095	.04	.84	.000
Error	211.757	95	2.229			
Total	3247.000	99				
Corrected Total	246.747	98				

a R Squared = .142 (Adjusted R Squared = .115)

Table 3. Regression analysis results for the effects of complexity, decision aid support, self-efficacy, and cognitive load on mental model accuracy

Predictors	B	Std. Error	β	t	Sig.	R^2 Total (Adjusted R^2)
(Constant)	.539	.181		2.97	.004	.271 (.238)
Complexity	.104	.052	.196	2.00	.048	
Decision Aid	-.118	.049	-.22	2.39	.019	
Self-efficacy	.08	.017	.464	4.68	.000	
Cognitive Load	.021	.017	.124	1.23	.22	

Dependent Variable: Mental Model Accuracy

Table 4. Two-way ANOVA results for the effects of decision aid support and task complexity on mental model accuracy

Dependent Variable: Mental Model Accuracy

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Corrected Model	.898(a)	3	.299	4.66	.004	.130
Intercept	107.840	1	107.840	1684.62	.000	.947
Complexity	.110	1	.110	1.71	.19	.018
Decision Aid	.540	1	.540	8.43	.005	.082
Complexity * Decision Aid	.251	1	.251	3.92	.05	.040
Error	6.017	94	.064			
Total	116.859	98				
Corrected Total	6.915	97				

a R Squared = .130 (Adjusted R Squared = .102)

Table 5. MANOVA results for between-subjects effects of decision aid support and task complexity on performance

Dependent Variable: Performance - trial blocks 1-6

Source	Type IV Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power(a)
Intercept	57.478	1	57.478	117.024	.000	.571	1.000
Complexity2	30.277	1	30.277	61.643	.000	.412	1.000
DecisionAid2	.299	1	.299	.609	.437	.007	.121
Complexity2 * DecisionAid2	.000	1	.000	.001	.978	.000	.050
Error	43.222	88	.491				

a Computed using alpha = .05

Table 6. General Linear Model results for the immediate-transfer phase (trials 4, 5, 6) of performance

Dependent Variables: Performance 4th, 5th, and 6th Trial Blocks

Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**	**Partial η2**
Intercept	.603	1	.603	2.52	.12	.029
Self Efficacy	.753	1	.753	3.15	.08	.036
Mental Model Accuracy	2.083	1	2.083	8.70	.004	.093
Cognitive Load	.152	1	.152	.64	.43	.007
Complexity	9.568	1	9.568	39.98	.000	.320
Decision Aid	.012	1	.012	.05	.82	.001
Complexity * Decision Aid	.416	1	.416	1.74	.19	.020
Error	20.343	85	.239			

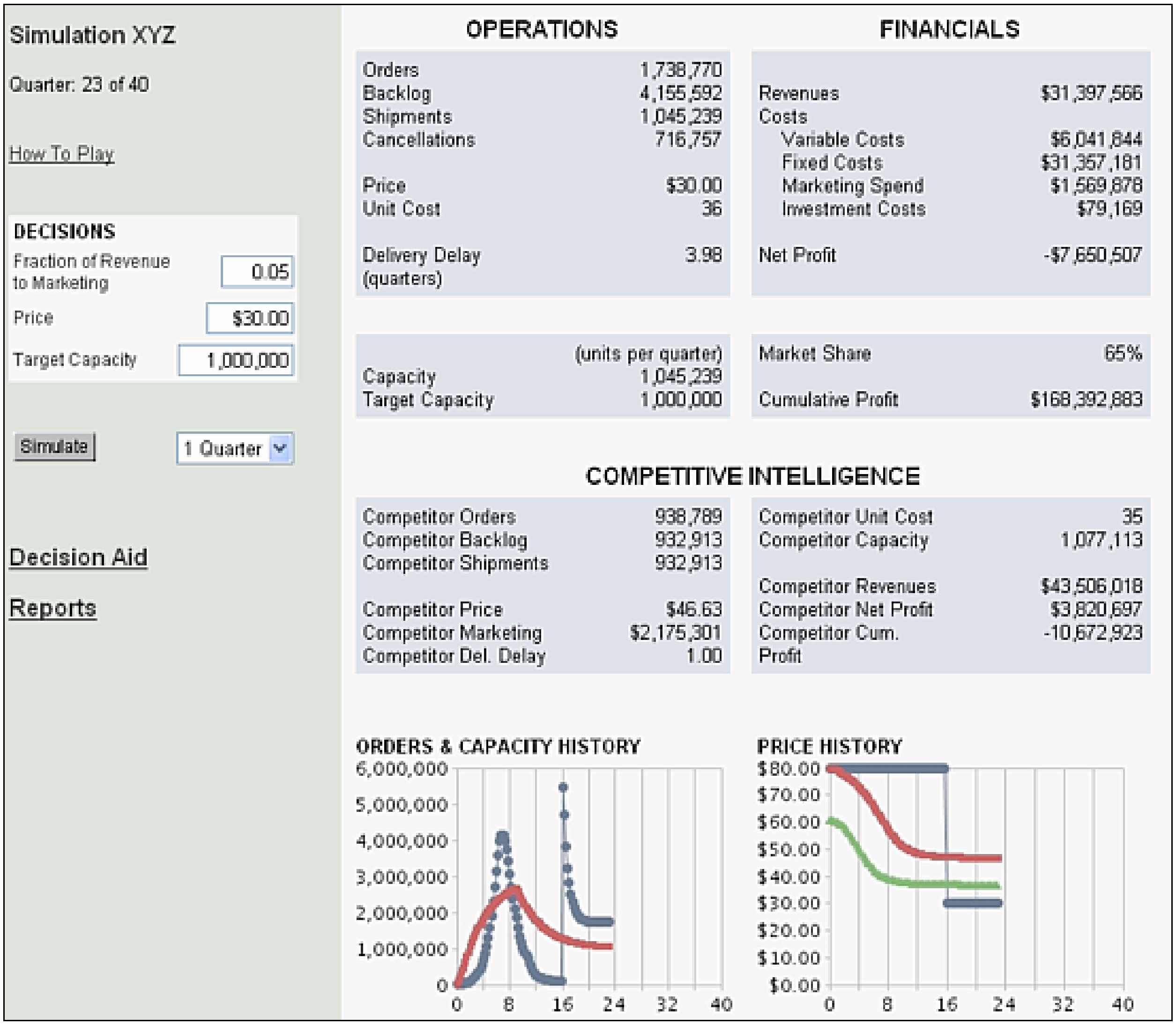


Figure 1. Screenshot of the High Complexity version of the simulation
Note: The three decision variables ‘price’, ‘target capacity’ and ‘Fraction of Revenue to Marketing are entered on the left side of the screen. For those in the decision aid condition, the decision aid is accessed via the ‘Decision Aid’ link on the left side of the screen.

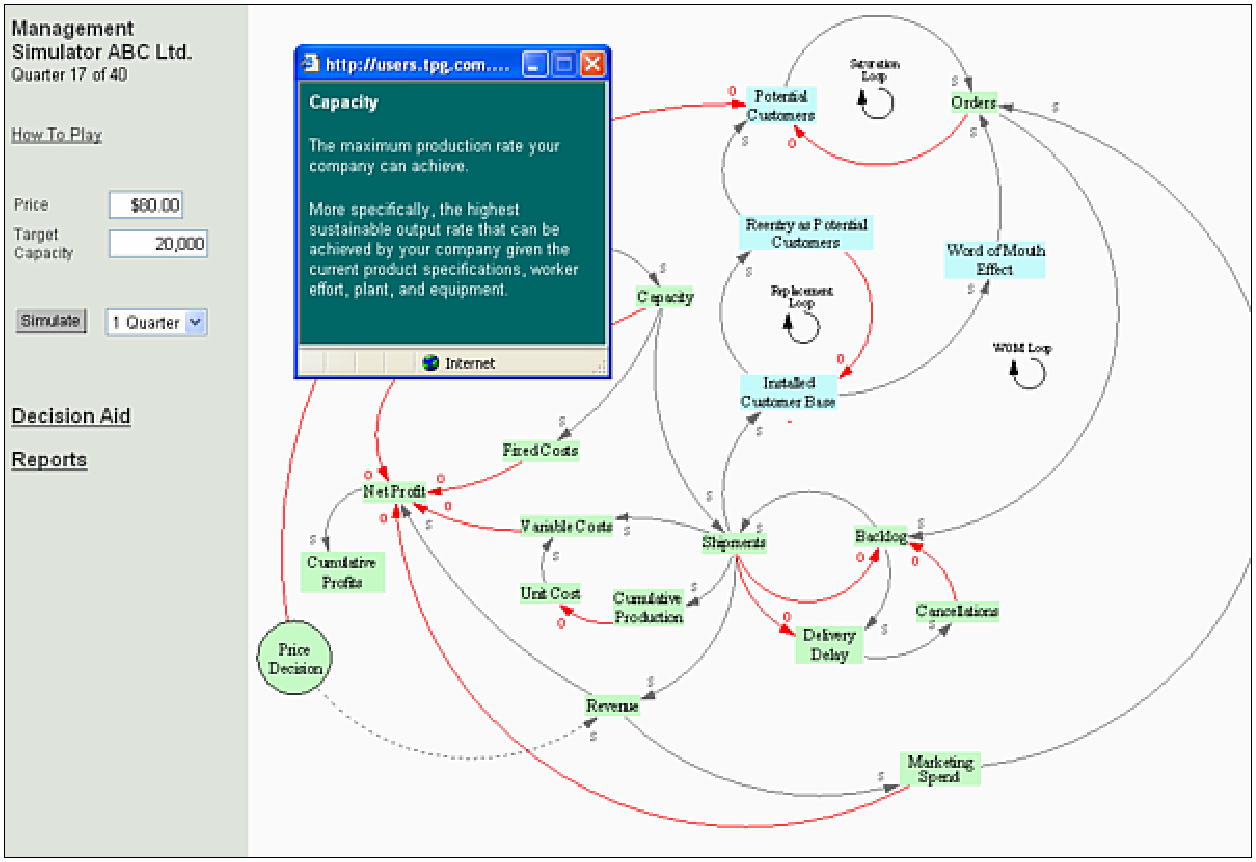


Figure 2. Screenshot of the Low Complexity 'Casual Map' Decision Aid**.** Clicking on any of the variables or arrows brings up an explanatory pop-up window.

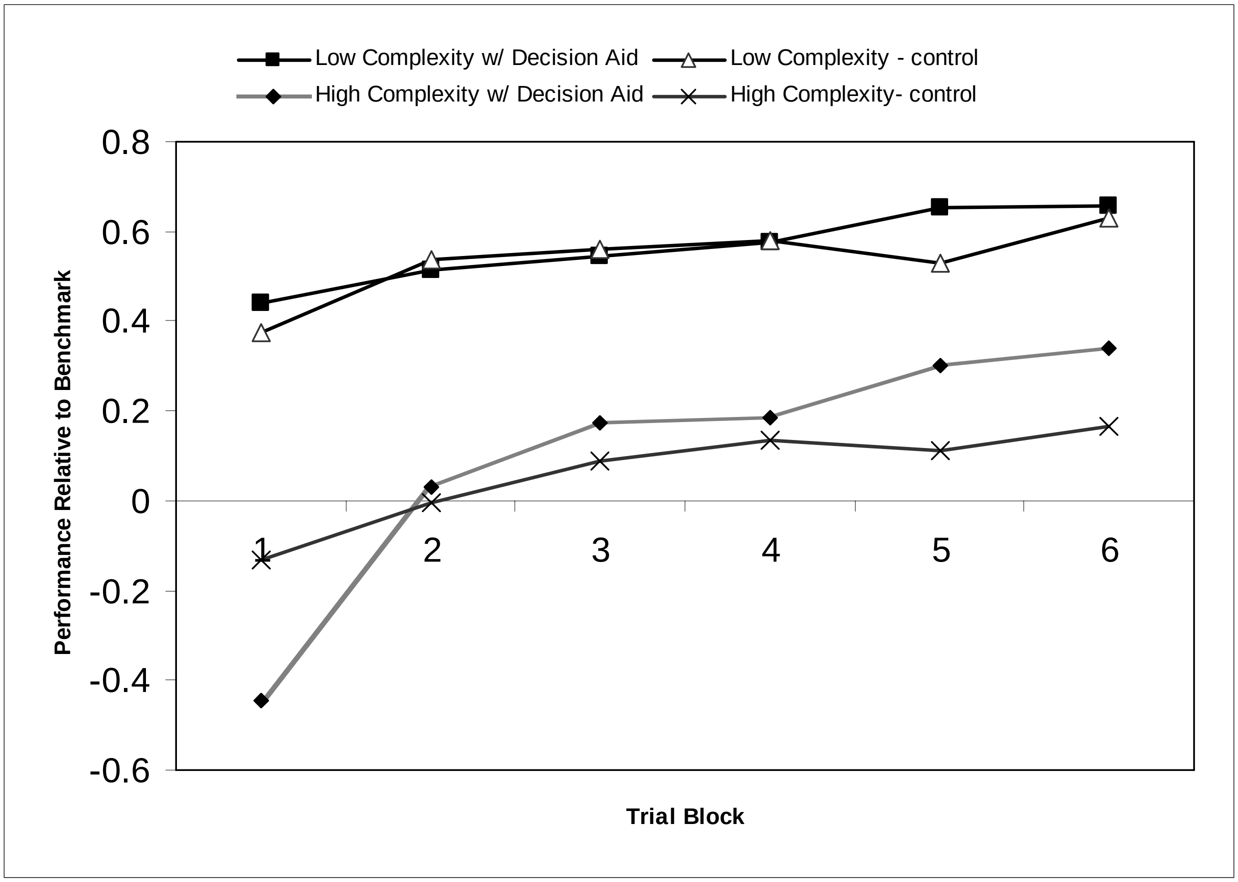


Figure 3. Mean performance relative to the benchmark for the four groups across the learning and immediate-transfer trial phases