Impacts of Dynamic Decision Making and Policy Development Modes on the Causal Understanding of Management Flight Simulators

Dalton E.M. McCormack[1] and David N. Ford[2]
Department of Information Science
University of Bergen
Norway

Abstract

System dynamics contends that understanding the causal structure of systems is required to successfully manage their behavior. However the best means of improving causal understanding is not well understood. Management Flight Simulators (MFS) have been used to study the effect of simulated environments on learning and management performance. These tools have been used almost exclusively to simulate dynamic decision making environments in which subjects make many decisions over the simulated time horizon. However recent research suggests that policy development MFS environments can improve causal understanding better than dynamic decision making environments. Policy development environment uses one continuous simulation through the time horizon as opposed to a decision making environment which involves stepwise simulation. We use an experiment to test this hypothesis by having subjects manage a small ecosystem in either a dynamic decision making or policy development mode. We quantitatively measure causal understanding with performance and an on-line questionnaire about the causal relationships in the managed system. Although the hypothesis is not supported a disaggregation of causal understanding into causal link and causal delay understanding suggests that a policy development mode helps to improve polarity causal understanding more than a decision making mode.

Introduction

A Management Flight Simulator (MFS) which is also known as a "Microworld" is a formal model in which decision makers can learn, refresh decision-making skills and play. Management Flight Simulators have two primarily uses. They help managers understand the interconnected nature of systems and the consequences of their actions. Designing effective MFS is an important issue for system dynamicists because they are important learning and training tools and central to the application of system dynamics in many domains. Ineffective MFS can fail to produce the benefits of system dynamics, damage the methodology's reputation and limit opportunities for expanding the methodology. Several issues must be addressed in the design of MFS. For example decision concerning the types of information to be included and the transparency of the underlying model structure must be made. Should the MFS's interface be a "black box" or "white box"? A "black box" MFS is an MFS with the underlying structure of model hidden from the user interface. A "white box" MFS is an MFS with the complete structure of the model shown in the user interface.

The literature suggests two methods of implementing MFS based on the mode of operation of the user (Langley and Morecroft, 1996): "stepwise simulation", sometimes called "gaming" and "continuous simulation", sometimes called "simulation". Based on the context of our

[1] Master's degree candidate, System Dynamics Program. University of Bergen, Norway. <Daltonm@ifi.uib.no>
[2] Associate Professor, System Dynamics Program. University of Bergen, Norway. <David.Ford@ifi.uib.no>

research we refer to the first mode of operation as a dynamic decision-making mode and the second as a policy development mode. In a dynamic decision-making mode subjects design and implement system management decisions repeatedly at regular intervals throughout each simulation of the system over the time horizon. In a policy development mode subjects design and implement their policies once at the start of the simulation which runs for the entire time horizon (Simons, 1990). Which of these two modes of operation is more effective and what impact do they have on learning in simulated dynamic environments? As a step in addressing this issue we investigate the effect of using either a dynamic decision-making or policy development mode on causal understanding.

Problem Description

Managing causal dynamic systems is often difficult. Designing tools that can help improve decision making and policy development is a primary focus of system dynamics. Management Flight Simulators help mangers learn and improve the performance of their systems by changing their mental models in a way that improves their understanding of structure-behavior relationships. People presumably transfer this greater understanding of the simulated system to other similar dynamic environments. As a part of the work to develop effective MFS they have been used as experimental tools to study decision-making processes of humans (Sterman, 1989), transfer of knowledge, system understanding across domains and the effect of different feedback types on performance (Bakken 1993, Abdel-Hamid and Sengupta 1993). Doyle, Radzicki and Trees (1996) found evidence of a positive relationship between using MFS and changes in mental models. Despite this progress research to date has been done only in dynamic decision-making contexts. Using a MFS in this mode requires time to gain understanding of the structure-behavior relationships of the system (Simons, 1990). Simons (1990), Langley and Morecroft (1996) suggest that changing from a decision-making mode of operation to a policy development mode will lead to improved learning in dynamic environments. Although this has been suggested as a way of improving learning in dynamic environments, little work has been done to test this hypothesis. Most of the MFS-based research has focused on using MFS as a tool to investigate other research questions such as the decision-making processes of humans in dynamic environments (Sterman, 1989). Our knowledge of how to build MFS which are effective learning tools is currently limited. The existing work has not adequately addressed operation mode which we believe is important in the design of effective MFS. Here we investigate the relationship between the operation mode used in MFS and causal understanding.

Research Hypothesis

We hypothesize that MFS based on a policy development mode of operation improve causal understanding of the simulated model more than those based on decision making. In a policy development mode subjects can experiment more by running more simulations within a given time span than in a decision making mode. We suspect that this allows policy development subjects to improve their causal understanding faster. In this work we limit causal understanding to the understanding of the relationships between pairs of variables in the simulated model such as those described with causal loop diagrams (see Appendix for example from the experimental model) and ignore more aggregate (but also important) forms of causal understanding such as feedback loops. Researchers have suggested such a relationship between the mode of operation and the effectiveness of MFS for learning (Simons 1990; Langley and Morecroft 1996) based on the reasoning that in a policy mode brings subjects closer to the model building process, which is required for greater insight about the real system (Langley and Morecroft, 1996; Sterman, 1994).

Our hypothesis is also suggested by the common problem experienced in the dynamic decision-making mode in which subjects frequently change their policy in response to outcome feedback in the form of system behavior without investigating the results of consistently implementing a single policy. This prevents retards their development of an understanding of the causal relationships of the system. Sterman (1994) argues for a structured form of experimentation in a virtual world (MFS) as a means of learning, of which causal understanding is a critical part. The additional structure imposed by the policy development mode may facilitate this kind of learning. Based on this reasoning we propose hypothesis

H1: The increase in subjects causal understanding in a policy mode will be greater than in the decision mode.

We measure causal understanding with system management performance and a questionnaire, as described later. Although our hypothesis that a policy development mode of operation in MFS increases causal understanding more than a decision-making mode has been suggested and is supported by the existing literature it has apparently never been rigorously tested.

Experimental Design, Tool and Procedure

We test the hypothesis with a true experiment (Campbell and Stanley, 1963) with differing treatments in which subjects where randomly divided into two groups: the Policy Development group and the Dynamic Decision-Making group and asked to manage a MFS of a small ecosystem. Bakken (1993) has shown that using an MFS that is based on a domain with which subjects are familiar can have an influence on causal understanding and performance. Therefore we selected a dynamic environment with a relatively simple context and in which minimal domain knowledge is required of subjects The simulated environment is based on a system dynamics model of the Kaibab Plateau model previously used to study how information structures impact performance in a policy development environment (Ford, 1997). Our version of this classic system dynamics predatory-prey model (Goodman,1974; Roberts, Andersen, Deal, Garet and Shaffer, 1983; Sterman, undated) dynamically models three species (deer predators, deer and grass) and provides managers with four control parameters: annual grass seeding, deer hunting, predator hunting and predator importation. We suspect that the bounded rationality[3] of subjects is often exceeded by the complexity of MFS, causing the loss of all treatment effects due to the overwhelming difficulty of the task. Therefore we used a relatively simple model of the ecosystem to not overwhelm treatment effects with system complexity. The user interface of the two operation modes were kept similar and simple, containing graphical and tabular information about some of the variables in the system in two frames which together fit legibly on a single page. However the system remains dynamically complex with a frequency of two oscillations over the forty year time region of the simulation.

Subjects in two locations were used in the experiment to control for educational and environmental biases. The subjects in Bergen, Norway where paid a hundred Norwegian Kroner (about 13 US Dollars) for participating and an additional hundred Kroner was awarded to the subject with the highest average score in the causal questionnaire in each group. Subjects in Worcester, Massachusetts were awarded 50.00 US Dollars for the best average score, 25.00 for second best and 10.00 US Dollars for third best. The experimental design and measuring parameters are shown in Table.1.

[3] Bounded rationality describes the limitations on human cognitive processing as central to understanding human behavior and performance (Simon,1974).

Group	Pre-Test	Treatment	Post-Test
Policy	Causal Questionnaire	Policy Development Tool.	Causal Questionnaire & Performance
Decision	Causal Questionnaire	Decision Making Tool	Causal Questionnaire & Performance

Table.1 Experimental Design and Parameter Measures

The Experimental Tool

The Kaibab plateau MFS consists of a system dynamics model and a user interface. The system dynamics model consists of three species stocks (predators, deer and grass), two first-order delays and the connecting rates and auxiliaries. Ford (1997) provides a description and complete documentation equation listing. Figure.1 shows the user interface of the Kaibab Plateau MFS. Subjects can see how the deer population unfolds over time in graphical as well as in tabular form (Figure.1). They can also see a record of their past decisions or policies on the table. Subjects also have a numerical performance indicator as shown in Figure 1 for a policy development performance. Larger performance numbers indicate better performance.

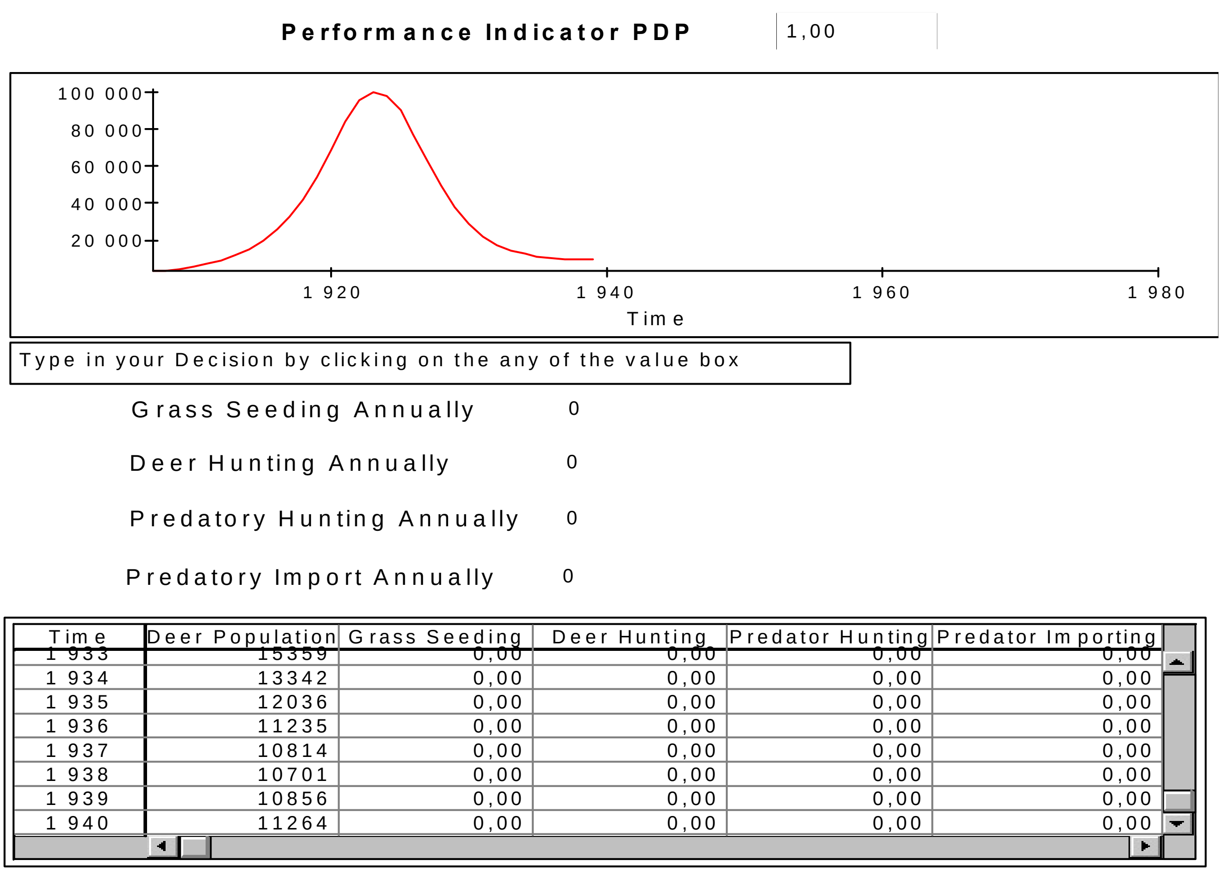

Figure. 1 The Kaibab Plateau MFS Interface

Two quantitative measurements were collected to measure system understanding: management performance and answers to a questionnaire. Management performance was measured by how closely the subject maintained the deer population to the goal 30,000 over the time period with the aggregate variance over time. The causal understanding questionnaire (Figure 2) which was based on a similar experiment by Bakken (1993) contained questions about the causal relationships in the model. For example the answer to the question " An *increase* in Grass Eaten by Deer leads to ... in Grass" (Figure 2) is "Immediate Decrease" since an increase in "Grass Eaten by Deer" will lead to an immediate decrease in the amount of grass in the plateau. The multiple choice questionnaire was computer based and contained 25 questions on causal relationships between variables in the system. The order of the questions was randomized and each question appeared on the monitor for a period of 25 seconds, followed by a new question. Each subject saw all 25 questions.

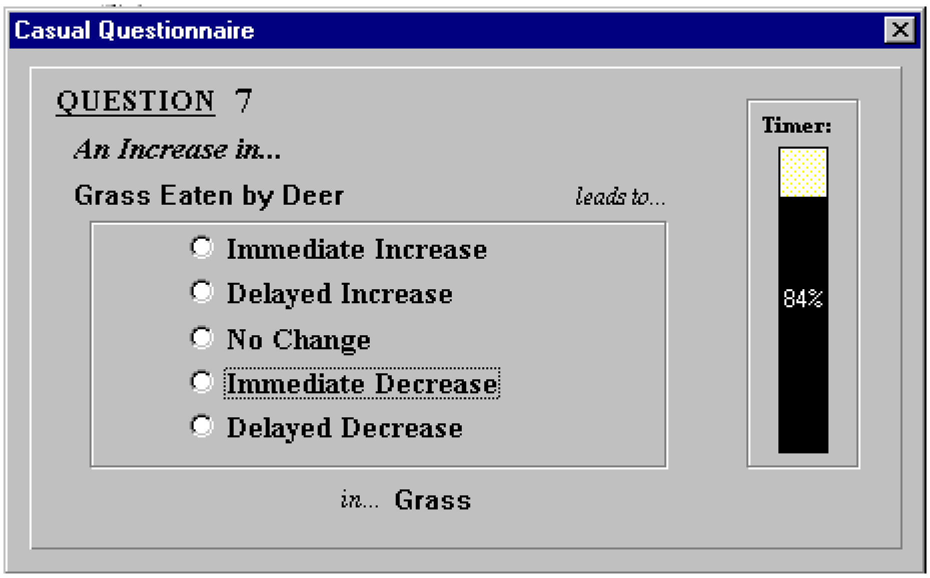

Figure. 2 Example question in the Causal Questionnaire

Experimental Procedure

Both policy and decision groups were given a description of the Kaibab Plateau based on Goodman (1974) and the questionnaire. The pre-test questionnaire was exactly the same for each group. Subjects in both groups were then asked to play a game in which they were required to control the deer population at a level of 30,000 for a period of 40 years. They used the four control parameters i.e. grass seeding, deer hunting, predator hunting and predator importation to do this for 20 minutes and saved each trial on a diskette. In each trial subjects type in their decision or policy on the "input" boxes and then click on the "run" button (not shown in figure 1). The simulation starts at zero and then stops at the year 1940. At this point subjects start to implement their decisions or policies. The difference between the two groups (the treatment) was that subjects in the decision group where required to make decisions annually (stepwise) throughout the forty year time horizon whereas those in the policy group where required to develop policies and to implement the policies at the start of each simulation. In other words subjects in the policy group were unable to stop the simulation after it has started.

At the end of the 20 minute session all subjects completed a post-test questionnaire. The post-test questionnaire was the same as the pre-test except that the order of the questions was changed at random. After the post-test questionnaire subjects were asked to fill out another questionnaire to control for variables such as age, educational background and experience in system dynamics and environmental studies. It also contained questions about the policy or decision rule the subject implemented while playing the decision or policy tool game.

Results and Analysis

The sample consists of 38 undergraduate and graduate students from the University of Bergen in Bergen, Norway and Worcester Polytechnic Institute in Worcester MA, USA. Eighteen subjects participated from the University of Bergen and twenty from Worcester Polytechnic Institute. The average age of subjects was 25 years, with a male/female ratio of 2 to 1. Some of the students had some experience in system dynamics although this was not a requirement for the experiment. Questionnaire performance was measured by assigning points as follows: 1 point was awarded for each question in which the answer was correct, 0.5 point if the answer to the question included either the correct polarity or delay relationship but not both and 0 point if the answer to the question contained no correct portions. The pre-test and post-test questionnaire scores of each subject together with their decision or policy development performance for each trail were recorded. The difference between the post-test and the pre-test causal understanding is denoted as the Causal Understanding Improvement (CUI).

Statistical analyses of the questionnaire scores (CUI) and performance scores using two sample t-tests do not support the hypothesis H1: that policy making mode increases causal understanding more than decision making mode. We found no correlation between policy or decision making performance and causal understanding, supporting a similar hypothesis tested by Bakken (1993) in a dynamic decision-making mode. This result contrasts sharply with the intuitive belief that improved causal understanding improves performance. Our results suggest that other necessary requirements for improved performance were not provided in our or Bakken's experiment.

In an effort to understand how operation mode impacts system understanding we disaggregated the causal understanding questionnaire data in two dimensions: degree of difficulty and type of understanding. There were three levels of difficulty: Easy, Medium and Difficult. The nine "Easy" questions involved one or two causal links between the variables (see Appendix). The twelve "Medium" questions had three to five causal links and the four "Difficult" questions had six or more causal links between the two variables. Based on this disaggregation we tested the following the following hypotheses:

H2: Subjects in the Policy Group improve their understanding of the simple causal relationships more than subjects in the Decision Group.

H3: Subjects in the Policy Group improve their understanding of the medium causal relationships more than subjects in the Decision Group.

H4: Subjects in the Policy Group improve their understanding of the difficult causal relationships more than subjects in the Decision Group.

Our results support hypothesis H2 that subjects in the Policy Group improve their understanding of the simple causal relationships more than subjects in the Decision Group ($p < 0.05$). Our results do not support hypotheses H3 and H4.

We also divided the causal understanding questions into two types of causal understanding: links with immediate impacts and links with delayed impacts. The nine questions with "Immediate Increase or Decrease" as the correct answer are referred to as "links with immediate impacts" questions and the sixteen questions with "Delayed Increase or Decrease" as the correct answer are referred to as "links with delayed impact" questions. Although the option "No change" was given as an answer alternative in the questionnaire interface, no questions had "No change" as the correct answer. Based on this disaggregation we tested the following the following hypotheses:

H5: Subjects in the Policy Group improve their understanding of the links with immediate impact more than subjects in the Decision Group.

H6: Subjects in the Policy Group improve their understanding of the links with delayed impact more than subjects in the Decision Group.

The results also support hypothesis H5 that there was more improvement in the immediate impact causal link understanding in the policy group than in the decision group ($p < 0.05$). However there was no significantly larger increase in the policy group in the understanding of links with delayed impacts, failing to support hypothesis H6.

Discussion

The main hypothesis was not supported by the results. This may be because the time spent using the MFS in both modes was too small (20 min) to cause a significant improvement in causal understanding or because other necessary learning factors were absent. Sterman (1994) suggests that bringing subjects closer to the modeling process can improve learning in dynamic environments. It's possible that a more sophisticated policy development tool in which subjects can build feedback loops into the system such as was used by Ford (1997) can generate improvements in causal understanding. Despite the failure to support the main hypothesis our results **do** indicate that the policy development mode improves system understanding more than the dynamic decision-making mode. The intersection of the two supported hypotheses (H2 and H5) are the simplest causal links with immediate causal impacts. This clearly indicates that the policy development mode improves system understanding in the least complex relationships of the system but could not generate improved system understanding beyond that of the dynamic decision-making mode concerning more complex relationships.

Conclusions

We used an experimental approach to test the hypothesis that the policy development mode impacts improvement in causal understanding more than the decision making mode and five other hypotheses. We measured causal understanding with an online questionnaire about causal relationships in the system and system management performance of a simulated dynamic ecosystem. Although our main hypothesis was not supported the support of two of our supplementary hypotheses indicate that operation mode does impact causal understanding differentially.

This work is an important step in understanding the effects of the operation mode on causal understanding. Additional research is needed to identify the necessary factors for improved causal understanding of complex as well as simple relationships and performance, which we suspect caused us to find support for only two of our six hypotheses. Further experiments are needed to test our hypotheses under these conditions and improve the measurement of causal understanding. It is important to also extend this line of research into other and more realistic simulated environments such as the "Boom and Bust" MFS (Paich and Sterman, 1993). Continued improvement of our understanding of the relationship between operation mode and causal understanding will allow the design and development of more effective management flight simulators for learning.

References

Bakken, Bent E (1993), Learning and Transfer of Understanding in Dynamic Decision Environments, *Ph.D. Dissertation Massachusetts Institute of Technology*.

Campbell, D. T. and Stanley, J. C. (1963) Experimental and Quasi-Experimental Designs for Research. Houghton Mifflin Co. Boston, MA.

Doyle, James K. and Ford, David N. (1997) Mental Models Concepts for System dynamics Research *Working Paper Reports In Information Science, University of Bergen.*

Doyle, James K., Radzicki, Michael J. and Trees, Scott W. Measuring the Effect of Systems Thinking Intervention on Mental Models, *Proceedings of the 1996 International System Dynamics Conference.*

Ford, David N. (1997) Structuring System Information for Improved Policy Development, *International Simulation and Gaming Association Conference July 7-10, 1997.*

Lane, David C.(1995) On a resurgence of management simulations and games, *Journal of the operational research society*, 46(1995), 604-625.

Langley, Paul A.(1996) Using Cognitive Feedback to improve Performance and Accelerate Individual Learning in a Simulated Oil Industry. *Working Paper under review at Management Science*

Langley, Paul A. and Morecroft, John D.W (1996) Learning from Microworld Environments: A Summary of the Research Issues, *Proceedings of the 1996 International System Dynamics Conference.*

Paich, M. and Sterman J.D (1993) Boom, Bust, and Failures to learn in experimental markets, Management Science, 39, 1439-1458.

Roberts, Nancy, Andersen, David, Deal, Ralph, Garet, Michael and Shaffer, William. (1983) Introduction to Computer Simulation, A System Dynamics Modeling Approach. Addison-Wesley Publishing Co. Reading, MA.

Sengupta, K and Abdel-Hamid, K.(1993) Alternative conceptions of feedback in dynamic decision environments: an experimental investigation. *Management science*, 39(1993), 411-428.

Simon H. (1974) How Big is a Chunk? *Science*, 183,482-488

Simons K.L.(1990) New Technologies in Simulation Games, *System Dynamics 1990 Conference Proceedings Vol III,* 1047-1059.

Sterman, John D.(1994) Learning in about complex systems. *System Dynamics Review* vol.10, nos.2-3 (Summer-Fall 1994): 291-330.

Sterman, John D.(1989) Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management science*, 35, No.3 (1989), 321-340.

Sterman, John D.(undated) Principles of Dynamic Systems I. *MIT, Sloan School of Management System Dynamics Group.*

Appendix: Causal Loop Diagram of the Kaibab Plateau Management Flight Simulator Model