Human–computer interface design can reduce misperceptions of feedback

Edward Howie, Sharleen Sy, Louisa Ford and Kim J. Vicente*

Edward Howie received an Industrial Engineering degree from the University of Toronto. He has over 11 years of experience with technology leaders including Delrina, MGI Software, and Archer Enterprises and is currently Director of Development Services at Cyberworld International Corporation

Sharleen Sy received an Industrial Engineering degree from the University of Toronto. She has over 11 years of experience with technology leaders including Delrina, MGI Software, and Alias Wavefront and is currently co-founder and Vice-President, Marketing at Cyberworld International Corporation

Louisa Ford is working towards her BASc degree in Industrial Engineering at the University of Toronto. She has worked as an undergraduate research assistant at the Cognitive Engineering Laboratory

Abstract

According to the misperceptions of feedback hypothesis, people make poor decisions because they do not heed the feedback dynamics in the environment, despite having perfect information about environmental structure. The thesis of this article is that dynamic decision-making performance can be improved by making the feedback structures of the environment more salient using human–computer interface design principles. An experiment was conducted in the context of Strategem-2, a management decision-making simulation used in previous research in this area. One group of participants used the original interface for Strategem-2, whereas another group used a new interface that was designed according to human–computer interface design principles. All participants were given a pre- and post-test to assess their knowledge of the simulation. The results reveal that the new interface leads to improved performance, and a greater improvement in knowledge of the microworld compared to the original interface. This finding is important because it suggests that misperceptions of feedback can be reduced through improved human–computer interface design. Copyright © 2000 John Wiley & Sons, Ltd.

Syst. Dyn. Rev. **16**, 151–171, (2000)

The global economy is becoming more and more complex as both the pace of change and the interconnectedness of actors increase in an unprecedented fashion (Davis and Meyer 1998). This increase in complexity has heightened the demands experienced by management decision makers. As a result, the study of dynamic decision-making in complex systems has become an increasingly important and timely research topic (for reviews see Dörner 1989/1996; Brehmer 1992; Sterman 1994). For the most part, the empirical findings obtained in this growing literature paint a pessimistic picture of human psychological capabilities. However, that research has under-emphasized the role that the computer interface can have on human performance. This issue is of great practical importance. As information technology has become more powerful and more affordable, the capability to present decision makers with new information displayed in new ways has improved tremendously. Moreover, research on human–computer interface design has shown that the way in which information is presented to people can have a strong impact—positive or negative—on their ability to cope with complexity (Woods 1991; Bennett and Flach 1992; Vicente and Rasmussen 1992; Vicente 1996).

The thesis of this article is that dynamic decision-making in complex systems can be improved by designing human–computer interfaces according to well-accepted human-factors design guidelines. This hypothesis was tested within

* Correspondence to Dr Kim J. Vicente, Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Road, Toronto, ON, Canada M5S 3G8. E-mail: benfica@mie.utoronto.ca

System Dynamics Review Vol. 16, No. 3, (Fall 2000): 151–171

Accepted January 2000

Kim J. Vicente is Professor of Mechanical & Industrial Engineering and of Biomaterials & Biomedical Engineering at the University of Toronto, and director of the Cognitive Engineering Laboratory. He is author of a recent book, *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-based Work*, published by Lawrence Erlbaum Associates

the context of a microworld simulation. However, if generalizable, the results have important practical applications for corporate managerial performance. More specifically, if human-factors design guidelines can improve human performance in a microworld, then those same guidelines could perhaps also be used to redesign interfaces to information system reports, thereby improving management decision-making in the turbulent, dynamic world outside of the research laboratory.

Background

Much of our understanding of dynamic decision-making in complex economic systems can be attributed to the extensive empirical work conducted by Sterman and his colleagues (Sterman 1987, 1989; Paich and Sterman 1993; Diehl and Sterman, 1995). One of the theoretical themes to emerge from this body of work is the misperceptions of feedback (MOF) hypothesis. From this perspective, people's poor performance in dynamic decision-making tasks in complex systems can be attributed to the fact that the ‘‘information needed for high performance is available but not heeded’’ (Paich and Sterman 1993, p. 1440). There could be several reasons for such misperceptions, each having different implications for our view of human performance in complex systems. The most pessimistic interpretation is that people have both the knowledge and information that is required to do well, but they nevertheless perform poorly because of some psychological frailty. This interpretation appears to be the one sometimes advocated by Sterman and colleagues. For example, Paich and Sterman (1993) stated: ‘‘even given perfect information and complete knowledge of the system structure people are not able to infer the resulting dynamics’’ (p. 1440; see also Sterman 1994; Diehl and Sterman 1995). On this reading, improved performance can only be achieved through automated decision support because human dynamic decision-making is bound to be poor because of ‘‘a fundamental bound on human rationality’’ (Diehl and Sterman 1995, p. 214). A less pessimistic interpretation of the MOF hypothesis is that people do not do well because their knowledge of system structure is less than perfect. In this case, poor performance is caused by lack of knowledge rather than some fundamental psychological limitation. On this reading, improved performance could perhaps be achieved through training by helping people to develop more accurate mental models of the system they are controlling (Diehl and Sterman 1995). A more optimistic interpretation of the MOF hypothesis is that people perform poorly because the information they need to do the task is not presented in the computer display that they have available to them. In this case,

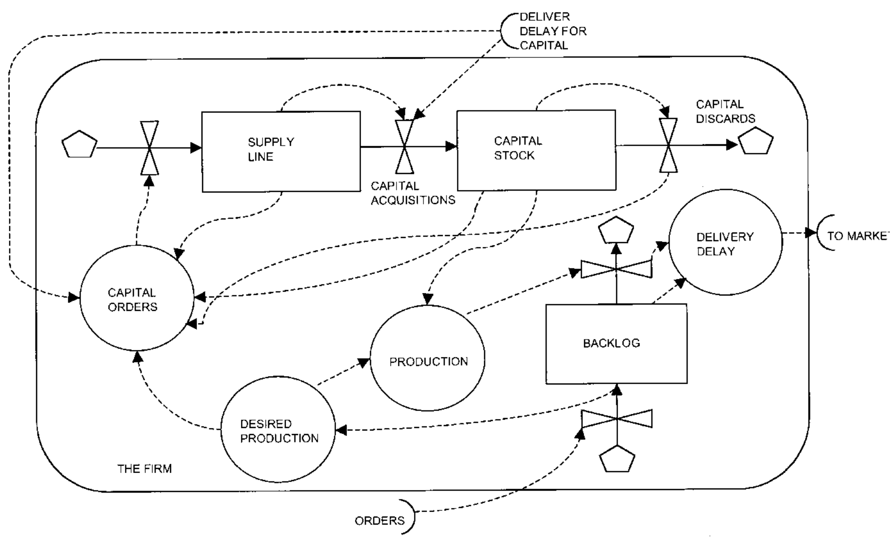


Fig. 1. The structure of the Strategem-2 simulation (adapted from Sterman, 1987)

poor performance is caused by a lack of information rather than a lack of knowledge or a fundamental psychological limitation. On this reading, improved managerial performance could be achieved through improved human–computer interface design. This possibility has been acknowledged by Diehl and Sterman (1995): ''Alternative displays and other decision aids may help overcome the misperceptions of feedback'' (p. 213). However, with the exception of a preliminary experiment conducted by Richardson and Rohrbaugh (1990), the possibility of reducing MOF via better interface design does not appear to have been systematically investigated. The research described in this article takes a step in filling this gap in the literature.

The Strategem-2 macroeconomic model developed by Sterman and Meadows (1985) was chosen as the testbed for this research. The basic structure of Strategem-2 is shown in Figure 1 and the following description is based on that provided by Sterman (1987). The simulation models the aggregate capital producing sector of an economy, which produces capital plant and equipment. The demand for the product is an externally determined variable. This demand must be satisfied by orders for production that accumulate in backlog until they are depleted by production. The amount produced at any one time is determined by capacity and capacity utilization. In turn, utilization is a nonlinear function of the current ratio of desired production to capacity. When the desired output is less than the capacity, then utilization is gradually cut back. When the desired output exceeds the capacity, then production is constrained by capacity. Capacity itself is determined by the capital/output ratio and the

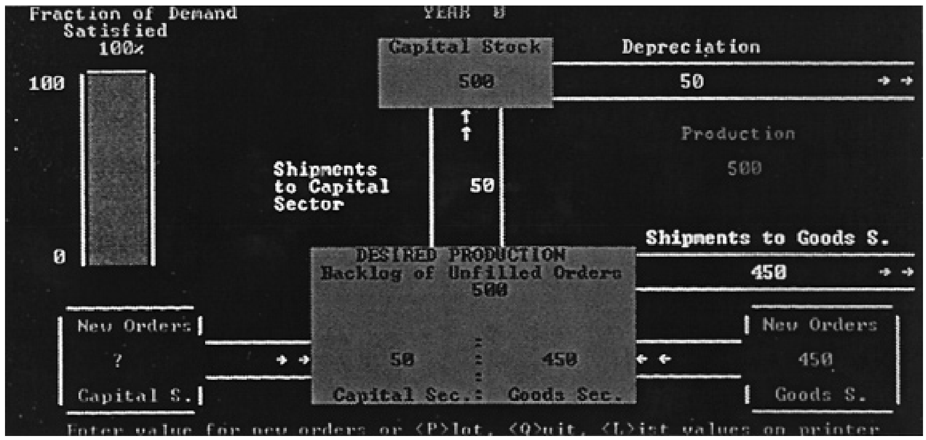


Fig. 2. The original interface for Strategem-2 used in Sterman's (1987) experiment

capital stock. Capital stock is increased by acquisitions and decreased by discards. The discard process is assumed to be exponential and the average lifetime to be constant. A key dynamic feature of the simulation is that orders for capital are only delivered after a delay representing the construction process. The company's orders for capital accumulate in the supply line until the construction process is completed and the capital enters the productive stock.

Strategem-2 was designed to operate in discrete time steps with an update interval of two years. Each update occurs after participants enter their decision. Players of the simulation only have control over one input, the amount of capital to order. Players must decide how much capital should be ordered so that the difference between the amount of product constructed and the amount of product being demanded is minimized. Thus, both undershooting and overshooting are undesirable. Note that, because of the lag in the system, orders placed in one period will only be received in the following update period. Performance with the simulation was measured as the average absolute deviation between production and demand, so that a low score indicates better performance.

Sterman (1987, 1989) investigated the ability of students and professionals with a business and economics background to make decisions in the context of Strategem-2. The interface that players used is shown in Figure 2. While the simulation is waiting for the player's input, nothing on the screen changes. Players input their order of capital for the current time interval in the box labeled ''New Orders–Capital Sector'' in the lower left corner of the display. The previous external demand input is shown in the box labeled ''New Orders–Goods Sector'' in the lower right corner of the display. Once the player enters their input, the simulation automatically updates all of the values shown in Figure 2. Color graphics and animation, not shown in Figure 2, were used to highlight the flows of orders, production and shipment.

In Sterman's (1987, 1989) experiment, each player was first given one practice trial. Then, they were given a second trial during which their performance was measured. Each trial lasted for 70 simulated years, with players entering an input every two years. The external demand for orders started off at 450 and, in year 4, rose to 500, where it remained for the duration of the trial. Thus, players were required to track one step input to the process. The average score was 432 (not including one outlier), which is considerably worse than the optimal score of 19. Moreover, most players generated an oscillatory pattern of repeatedly undershooting and overshooting the desired demand. Only 8% of the players were able to stabilize the system within the 70-year duration of the trial, despite the fact that the external demand remained constant at 500 since year 4.

Sterman (1989) introduced the MOF hypothesis to explain these results. Under this view, players controlled the simulation as if it was an open-loop system rather than a closed-loop positive-feedback system. Because the feedback loop was not taken into account, oscillations were generated and performance was poor. What makes this result particularly important is that, according to Sterman (1987), participants had ''perfect knowledge of the system structure and perfect information'' (p. 1587; see also Sterman 1994). Despite these apparently ideal conditions, performance was far from optimal.

We can re-examine these results and this interpretation in light of the three versions of the MOF described earlier. The claim of perfect knowledge is based on the fact that all players read a comprehensive description of the Strategem-2 game and were given a chance to ask questions before interacting with the simulation. However, just because a complete description of the simulation was provided does not mean that players were able to assimilate all of this knowledge. A test of players' knowledge was not administered, so it is not possible to determine if the assumption of perfect knowledge is reasonable. It is possible that at least some players did not have a complete understanding of the system structure, in which case their sub-optimal performance might be improved through training, rather than being an unalterable given due to a cognitive limitation.

Similarly, the claim of perfect information is based on the fact that the interface (see Figure 2) displayed all of the system variables on the screen at all times. However, in the human–computer interface design literature, a distinction is made between ''designing for data availability'' and ''designing for information extraction'' (Woods 1991). As the name suggests, designing for data availability makes all of the raw data that are required to perform a task available to a user, who then has to process those data cognitively to answer a higher-order question of interest. In contrast, designing for information extraction explicitly shows, not just the data, but also the higher-order relationships between data. Showing these relationships improves human performance by reducing the need to process low-level data using scarce cognitive resources. There is considerable empirical evidence in other settings showing that designing for information extraction leads to better performance than designing for data availability (Bennett and Flach 1992; Vicente 1996).

This finding has direct implications for the MOF hypothesis. If we compare the interface that players had to use when controlling Strategem-2 (see Figure 2) with the diagram showing the structure of the Strategem-2 dynamics (see Figure 1), we find differences between what was directly shown to players and what was really going on in the underlying simulation. Figure 2 looks much simpler than Figure 1. However, this simplicity is deceptive because it is Figure 1 that comprehensively describes the dynamics of Strategem-2, not Figure 2. Yet, players had to control the system with Figure 2 without ever having seen Figure 1. Of particular importance is the fact that the key system relationships are not perceptually specified in the interface in a salient manner. For example, the bottom of Figure 2 shows a linear flow of goods with two inputs and two outputs represented by large arrows. The positive feedback loop that is so critical to the system's dynamics is not explicitly shown. This fact becomes significant in light of the ''out of sight, out of mind'' phenomenon—people's natural tendency to treat a display as if it is the system being controlled, as opposed to a representation of that system (Fischoff *et al*. 1978). As a result, people tend to ignore relationships that are not explicitly represented in the display in a perceptually available manner. Therefore, it is possible that the sub-optimal performance of players interacting with Strategem-2 is at least partially a result of the characteristics of the interface that they were given, in which case performance could be improved through interface design rather than being solely limited by ''a fundamental bound on human rationality'' (Diehl and Sterman 1995, p. 214).

The experiment described in this article was designed to address these issues empirically. Players interacted with one of two interfaces to Strategem-2, the original one used by Sterman (1987, 1989) and a new interface that was designed to enhance information extraction. The hypothesis was that, as a result of making higher-order system relationships explicit in a perceptually salient fashion, performance would improve. In addition, players were given a pre- and post-test to evaluate the extent of their knowledge of the Strategem-2 simulation. The results of these tests allow us to determine if the assumption of perfect knowledge is accurate, and if players' knowledge improves with practice with each interface. The results of the experiment should provide us with a better idea of whether the MOF hypothesis should be interpreted as a cognitive frailty that cannot be remedied, or whether it is a limitation that can be reduced through better training or improved interface design. These findings will, in turn, have important practical implications for how corporate managerial performance in complex systems may be improved.

Method

Experimental design

A 2×2 mixed factorial design was adopted for the experiment with Interface (New and Old) as a between-participants variable and Trial (1 and 2) as a within-participants variable. Before and after the two trials, participants filled out a questionnaire that tested their knowledge of Strategem-2 (see Appendix).

Participants

Ten participants were randomly assigned to each interface group. All of the participants had some computer experience, were fluent in English and possessed a post-secondary degree. None of the participants had any experience with the Strategem-2 simulation. The average age of the participants was 29 years, with a range of 22 to 35 years.

Apparatus

The Old interface for Strategem-2, shown in Figure 2, was identical to the one used by Sterman (1987, 1989). It was programmed in BASIC and is available in a DOS version for PCs (Sterman and Meadows 1985). The New interface, shown in Figure 3, was created using Authorware for Windows, version 4.0 from Macromedia Corporation.[1] It was designed according to the following human–computer interaction principles (see Howie and Sy 1996 for a more detailed description):

- Wherever possible, take advantage of people's natural tendencies (e.g., reading from left to right and from top to bottom).
- Wherever possible, take advantage of people's existing knowledge (e.g., through the use of metaphors).

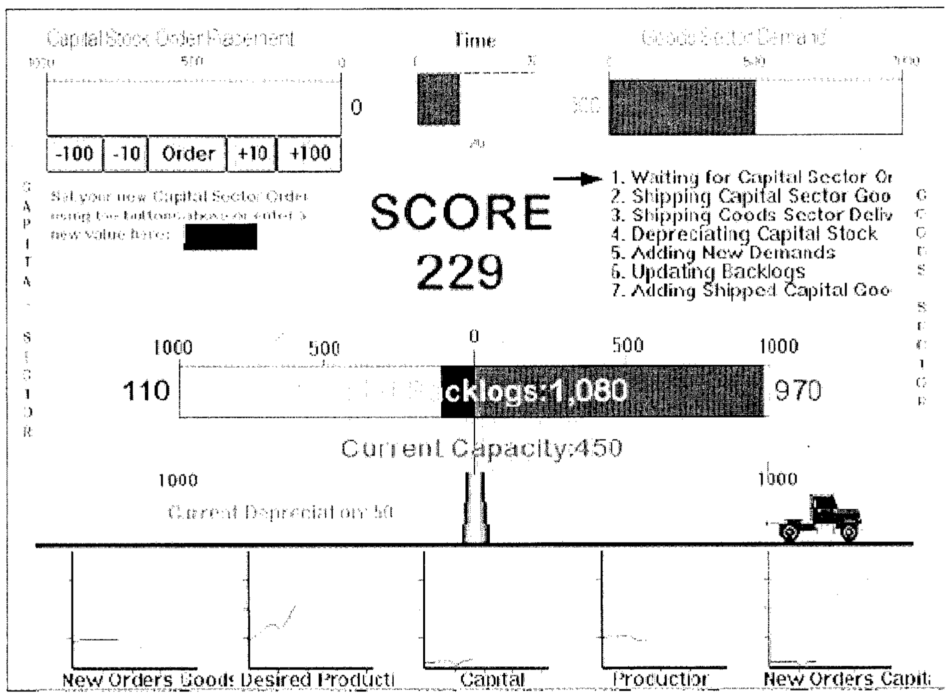


Fig. 3. The new interface for Strategem-2 used in this research

- Present information in a graphical, rather than just a text-based, fashion to exploit people's powerful pattern-recognition capabilities.
- Show, not just data, but also relationships between data so that people can develop a more accurate mental model of the simulation.

These design principles are based on well-accepted knowledge in the human factors literature (Woods 1991; Bennett and Flach 1992; Vicente and Rasmussen 1992).

In Figure 3, the horizontal bar in the top left of the screen shows the Capital Stock ordered by players. The horizontal bar in the top right of the screen shows the Goods Sector Demand, which is an exogenous input to the simulation. Between these two bars, the simulated time is shown. Directly below the time display, the player's current score is indicated. Immediately to the right of the score, a list of seven items is shown. These items represent the steps that the simulation goes through in updating its values. An arrow indicates the step that the simulation is currently at. For example, in Figure 3, the simulation is waiting for the Capital Sector Order from the player. This seven-step list provides an overview of the simulation's structure to make it easier for players to understand how the simulation works, and to know where they currently are in the update cycle.

In the middle of the display, the Total Backlog and the Current Capacity are shown as bars on a common scale that are supported by a metaphorical hydraulic lift, whose purpose will be described shortly. This design should allow people to see, and thus understand, the relationship between these two variables. In addition, the Total Backlog scale is divided into two parts to distinguish between the contribution made by the two sources. The contribution of the Capital Stock Order is shown to the left of the zero point in the Total Backlog scale, whereas the contribution of the Goods Sector Demand is shown to the right of the zero point in the Total Backlog scale. This coding should allow players to understand how the Capital Stock Order and the Goods Sector Demand, shown along the top of the interface in Figure 3, interact to determine the Total Backlog.

Animated graphics are also used to convey the dynamics associated with each of the steps in the seven-step list. For example, after players enter their order, the arrow moves to step 2 in the list, Shipping Capital Sector Goods. At this time, the hydraulic lift supporting the Total Backlog and the Current Capacity bars lowers the amount of Backlog that fits within the Current Capacity. This relationship can be seen graphically, because, if the Total Backlog bar is greater than the Current Capacity bar, not all of the Total Backlog will be lowered by the lift. After this step, the arrow in the seven-item list moves to step 3, Shipping Goods Sector Deliveries. At this time, the truck shown near the bottom right of Figure 3 drives up to the lowered bar and takes the goods to be shipped away from the Capacity bar. Again, the animation tries to show, in a concrete and intuitive way, abstract relationships that might otherwise be very difficult for players to apprehend. Similar animated graphics techniques are used to convey the dynamics associated with the remaining steps on the seven-item list (see Howie and Sy 1996 for more details).

It is important to note that the only difference between the Old and New interfaces was the presentation of information; the underlying dynamic model was identical.

Experimental tasks

Participants were required to control the Strategem-2 simulation by entering values for the capital stock order. The task was identical to that in Sterman's (1987, 1989) research. The update interval was two simulated years. The external demand for orders started off at 450 and, in year 4, rose to 500, where it remained for the duration of the trial. Each trial lasted for 70 simulated years.

Participants were told that their objective was to minimize their score by ordering an appropriate amount of capital stock.

The short questionnaire participants were required to fill out before and after they used the simulator evaluated participants' knowledge of Strategem-2 before they had any experience with the system and after they had some experience controlling the system with whichever interface they were given. Thus, the pre- and post-test comparison provides a means of evaluating the degree of learning that is induced by each interface.

Procedure

The experiment consisted of five phases: an introductory phase, during which the experimental protocol was explained to participants; a pre-test phase, during which participants were tested for their knowledge of Strategem-2; a practice phase, during which participants completed a practice trial with the simulation; a testing phase, in which participants completed two trials; and finally, a post-test phase, during which participants were tested once again for their knowledge of Strategem-2.

Introductory Phase. Each participant was seated in front of an IBM PC and was directed to read instructions for the game. The instructions were the same as those used by Sterman (1987, 1989) and provided a comprehensive description of the Strategem-2 simulation. After they had read the instructions, participants were permitted to ask questions.

Pre-test Phase. After the introductory session, participants were asked to fill out a short pre-test questionnaire to evaluate their knowledge of the system. The answers to all of the questions could be found in the instructions that participants had just finished reading, although participants were not allowed to consult those instructions when filling out the questionnaire. If participants had perfect knowledge of the system, then they should achieve a perfect score on the questionnaire. Feedback on the questionnaire score was not provided.

Practice Phase. In Sterman's (1987, 1989) experiment with Strategem-2, participants were provided one practice trial during which they could ask questions. Data from this practice trial were not analyzed. To make our results comparable, we followed the same procedure. Once the questionnaire was completed, participants completed one trial with Strategem-2 and were allowed to ask questions pertaining specifically to the instructions. Questions that related to game play strategies or that were not covered in the instructions were not answered.

Testing Phase. During this phase, participants completed two trials of 70 simulated years of duration each. The results reported below are from these two trials. After a participant entered values into the computer, a computerized sequence of events occurred as follows (Sterman, 1987):

1. The orders are received from the capital and the goods sector. The fields ''New Order–Capital Sector'' and ''New Order–Goods Sector'' are added to orders remaining from the prior periods.
2. Desired production is calculated for the current period and is found in the appropriate field on the screen.
3. Production is calculated for the current period and is displayed on screen. Production capacity is also calculated and displayed.
4. The newly produced capital is delivered to each sector.
5. Depreciation is calculated for the current period and displayed on screen.
6. Participant enters a new capital order and enters it in the ''New Orders–Capital Sector'' field, time is incremented by two years, and the cycle starts again.

Note that the formulae behind Strategem-2 were not changed in the New interface version; only a new display was created.

Post-test Phase. The same questionnaire was administered once again to each participant at the end of the experiment. These data help us determine if the type of interface has an impact on the knowledge that a participant acquired while controlling the system.

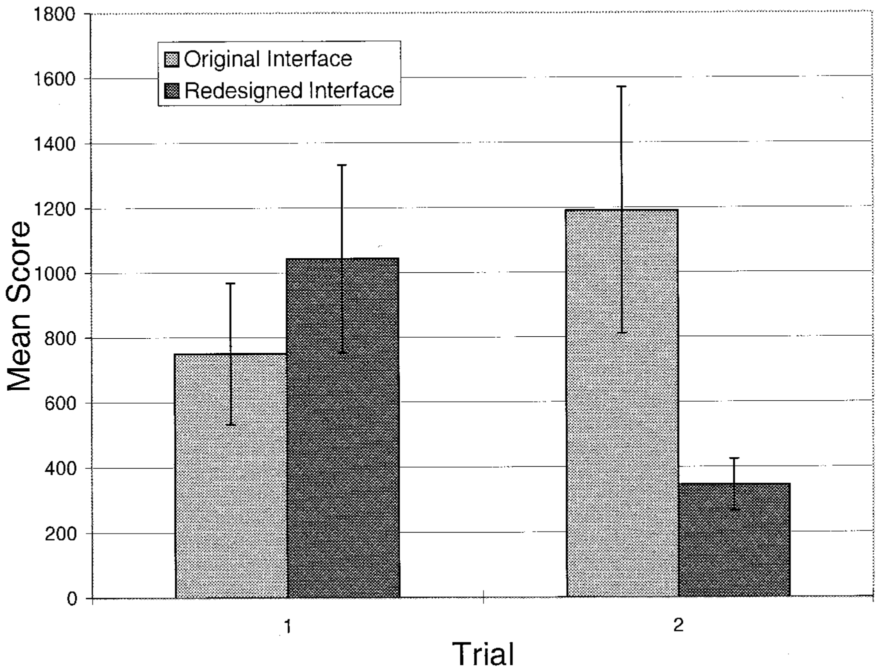
Dependent variables

Performance with Strategem-2 was measured using the same formula used by Sterman (1987, 1989), namely the average absolute deviation between production and demand. Thus, a low score indicates better performance. Performance on the questionnaire was measured as a percentage of the questions answered correctly.

Results

The performance results will be presented first. To control for the effects of prior knowledge or ability, the pre-test score was used as a blocking variable.

Fig. 4. The interaction between Interface and Trial for game score with standard error bars



Thus, the five participants with the highest questionnaire score in each interface group were classified as Knowledgeable, whereas the five participants with the lowest questionnaire score in each interface group were classified as Less Knowledgeable. A $2 \times 2 \times 2$ ANOVA was then conducted with Interface (New and Old), Trial (1 and 2), and Knowledge (Knowledgeable and Less Knowledgeable) as factors and simulation score as the dependent variable. As in Sterman (1987, 1989), data from the first practice trial were not included. As is to be expected, the main effect for Knowledge was significant ($F(1, 32) = 13.46$, $p < 0.0009$). The mean scores for Knowledgeable participants and Less Knowledgeable participants were 414.0 and 1250.6, respectively. In addition, the Interface × Trial interaction was significant ($F(1, 32) = 6.24$, $p < 0.0179$). No other effects were significant.

Figure 4 illustrates the crossover interaction between Interface and Trial. Recall that a lower score indicates better performance, and that the optimal score is 19. For the first trial, the Old interface group (mean score of 750.1) outperformed the New interface group (mean score of 1043.2). Additional experiments are needed to determine the cause of this finding with confidence, but one explanation seems particularly likely. The New interface tries to provide a comprehensive visualization of the microworld structure and dynamics. As a result, it is visually more complex than the Old interface because it is mirroring the complexity in the underlying simulation. It is possible that participants are initially not able to cope with this added visual complexity. Practice may be required to take advantage of the benefits that the New interface has to offer. This interpretation is consistent with the reversal of results obtained in the second trial. The New interface group exhibited a substantial improvement in performance, with a mean score of 344.9. In contrast, the performance of the Old interface group became worse on the second trial, with a mean score of 1190.9. Note also that the variability of the New interface was markedly decreased for the second trial, as evidenced by the narrow standard error bars. The standard deviation for the Old interface group was 1200.3, which notably exceeded the New interface group's standard deviation of 251.4. These results suggest a greater consistency in scores by the New interface group.

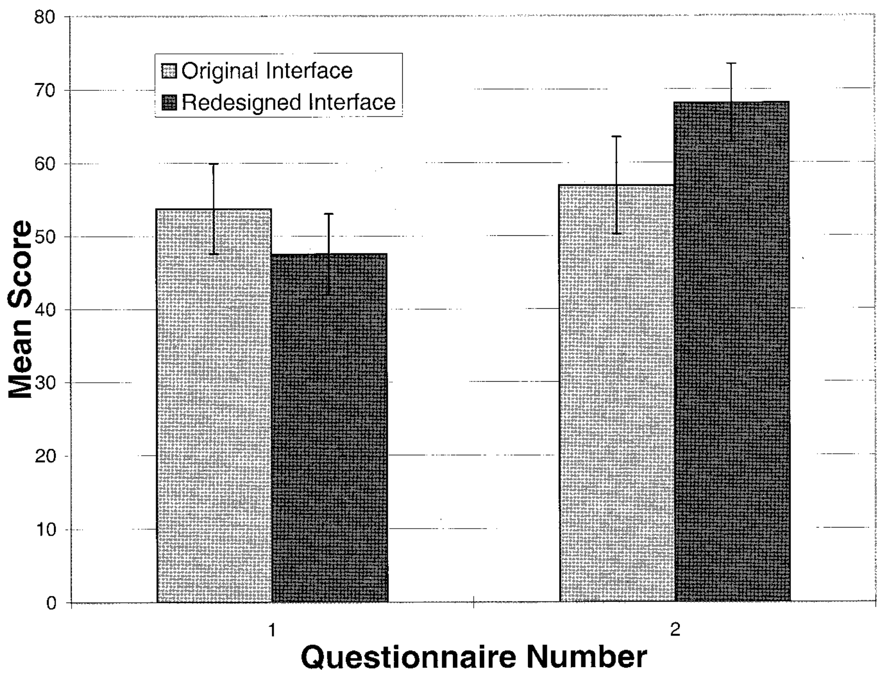
The questionnaire data were analyzed using a $2 \times 2 \times 2$ ANOVA with Interface (New and Old), Questionnaire Administration (Pre-test and Post-test), and Knowledge (Knowledgeable and Less Knowledgeable) as factors. Note that the blocking variable, Knowledge, was again based on the pre-test questionnaire scores. As expected, the main effect for Knowledge was significant ($F(1, 32) = 47.80$, $p < 0.0001$). The Administration effect was also significant ($F(1, 32) = 0.0054$, $p < 0.0054$), as was the Interface × Administration interaction ($F(1, 32) = 4.84$, $p < 0.0352$). No other effects achieved significance.

Figure 5 shows the main effect for Administration as well as the Interface × Administration Interaction. Not surprisingly, the mean score on the post-test was greater than on the pre-test (means of 62.5 and 50.6, respectively). More interesting is the Interface × Administration interaction. Before the participants had any experience with Strategem-2, the Old interface group exhibited a slightly higher pre-test score. The mean scores of the Old and New interface groups were 53.8 and 47.5, respectively. However, after experience with Strategem-2, a reversal was observed. The Old interface group exhibited only a marginal improvement (mean of 56.9), whereas the New interface group exhibited a much larger improvement (mean of 68.1), surpassing the Old group. These results show that practice has a larger beneficial performance with the New than with the Old interface. Note also that the means for both groups were well below the perfect knowledge level of 100%.

Discussion

What do these results teach us about the MOF hypothesis? First, the results clearly show that, at least for the case of Strategem-2, performance is limited by imperfect knowledge. Although participants read a set of instructions that

Fig. 5. The interaction between Interface and Administration for questionnaire score with standard error bars



provided a comprehensive description of the Strategem-2 simulation, they were not able to assimilate and retain all of this knowledge. As a result, their questionnaire performance was well below perfect, even after experience with controlling the system. These findings suggest that training could improve performance by reducing MOF.

Second, the results also clearly show that human–computer interface design can reduce MOF. After some practice, participants in the New interface group were considerably faster and more consistent than those in the Old interface group. Moreover, the New interface group learned more about the system than did the Old interface group, as evidenced by the greater improvements in questionnaire scores. These findings show that MOF are, in part, due to inferior interface design rather than just cognitive frailty. Redesigning human–computer interfaces can enhance both dynamic decision-making performance and system understanding.

Third, the results also show that, even with the New interface, performance was not as good as it could be. The optimal performance level of 19 was not achieved by any participant. Similarly, no participant achieved a perfect level of 100% on the questionnaire, even after some experience with the system. Thus, it seems that, with the modest level of practice provided in this experiment, there is still room for improvement. Our findings do not allow us to determine if more practice would lead to exceptional performance or whether psychological limitations impose a sub-optimal limit on performance.

Implications

What lessons does this research have for the system dynamics community? Three points stand out. First, it is important that participants be provided with some practice in microworld research. As shown in Figures 4 and 5, the impact of some independent variables (e.g. interface design) may only be observed with experience. Second, it is also important that participants' knowledge be tested directly if claims are to be made about knowledge or mental models. As shown in Figure 5, providing participants with a comprehensive description of how a microworld functions may be insufficient to ensure that people have perfect knowledge. Finally, and most importantly of all for the purposes of this article, these results open the prospect that corporate managerial performance could be improved by better human–computer interface design. There is an extensive literature consisting of human-factors design guidelines that have been demonstrated to improve human performance in other dynamic decision-making contexts (e.g., Woods, 1991; Bennett and Flach, 1992; Vicente and Rasmussen 1992; Wickens 1992). System dynamics researchers and practitioners can take advantage of these guidelines to redesign the human–computer interfaces of existing information system reports with the intent of improving performance and understanding as well.

Limitations

Although these results show that human–computer interface design can reduce misperceptions of feedback, the experiment was only conducted with one microworld simulation. To assess the generalizability of this finding, it is important to try to replicate the result in other microworlds. Perhaps even more importantly, it is also important to try to replicate the result under more representative conditions to demonstrate that corporate managerial performance can in fact be improved by human-factors design guidelines.

Conclusions

In the economy of the future, dynamic decision-making in complex systems will become more and more important. Given the marked improvement in the capabilities of information technology, it is economically feasible to present decision makers with new information formatted in new ways (e.g., redesigned IS reports). The empirical evidence presented in this article shows that human–computer interface design techniques can be used to reduce the difficulties people have in dealing with complex systems. At the same time, we have shed some new empirical light on the MOF hypothesis. Rather than being solely caused by an unalterable cognitive limitation, it appears that MOF are caused in part by lack of knowledge and in part by deficient interfaces. Perhaps by focusing more attention on training and especially improved interface design, we can help people make better decisions in the face of complexity.

Acknowledgements

This research was sponsored by research and equipment grants from the Natural Sciences and Engineering Research Council of Canada. We would like to thank John Rohrbaugh, John D. Sterman, and Gerard Torenvliet for their help with this research, and Professor Vennix and the two reviewers for their constructive and thorough comments.

Appendix
Questionnaire for Strategem-2

Subject Name:

1) Is there depreciation on goods shipped to the goods sector?
   Y/N

2) What are the sector(s) in the economy?
   a) Depreciation
   b) Goods
   c) Capital
   d) A and C but not B
   e) B and C but not A

3) Of which sectors (use question 2's options) do you have control over?
   Answer:

4) If the capital sector's demand increases and any other demands stay the same, the subsequent overall backlog:

a) Increases
b) Decreases
c) Does not change
d) The capital sector's demand is irrelevant to the backlog

5) The typical wave the Kondratiev long wave depicts is:
a) Exponential
b) Logarithmic
c) Cyclical (Damped Oscillations)
d) Cyclical (Amplified Oscillation)

6) Backlog consists of:
a) Capital sector orders
b) Capital goods depreciation
c) Goods sector orders
d) A and C but not B
e) None of the above

7) The factor(s) which contribute to the score consists of:
a) Total overproduction
b) Total underproduction
c) Total demand satisfied
d) A and B but not C
e) None of the above

8) To get a better score, you should:
a) Maximize production
b) Minimize over and underproduction
c) Minimize overproduction
d) A and B but not C
e) None of the above

9) How does the current capacity increase?
a) Capital sector shipments
b) Goods sector shipments
c) Depreciation
d) A and B but not C
e) None of the above

10) How does current capacity decrease?

a) Capital sector shipments
b) Goods sector shipments
c) Depreciation
d) A and B but not C
e) None of the above

11) Depreciation consists of:
a) The consumption of capital goods by the goods sector
b) Lost orders in transit to the goods sector
c) Capital order produced for the capital sector's consumption
d) Reduction in current capacity from wear and tear
e) None of the above

12) Does depreciation reduce the production capacity of the capital sector? Y/N

13) The capital sector:
a) Consumes the produced goods
b) Produces goods to be consumed by the goods sector
c) Consumes the depreciated material
d) A and C but not B
e) B and C but not A

14) Goods leave the production system via:
a) Shipment to the goods sector
b) Shipment to the capital sector
c) Depreciation
d) A and C but not B
e) None of the above

15) If the capital sector is running at full capacity, and the goods sector demand increases, in the next period meeting the increase in demand will cause:
a) The capital sector demand to increase
b) Underproduction
c) A backlog in the goods sector shipments
d) All of the above
e) None of the above

16) If the capital sector is larger than the production capacity of the capital

sector, will the goods sector's demand be fully satisfied in the next period? Y/N

Match an answer to the following questions. Note there are more answers than there are questions.

Questions:

1. What is represented by the field new orders–capital stock?
   Answer:
2. What is represented by the field New Orders–Goods Sector?
   Answer:
3. How do you obtain desired production?
   Answer:
4. How does the simulation calculate the current capacity?
   Answer:
5. What is represented by the capacity field?
   Answer:
6. What factors are involved in calculating fraction of demand satisfied?
   Answer:
7. What factors are involved in computing your score?
   Answer:
8. How is the current capacity proportioned?
   Answer:

Answers:

a) Ratio of capital sector backlog to goods sector backlog
b) Add capital sector backlog plus goods sector backlog
c) Cumulative total of previous year's over/under production divided by current simulation
d) Goods sector demand
e) Current capacity and total backlog
f) Capital sector demand
g) Total production capability for capital sector
h) Depreciate capacity, then add capital sector shipments to current capacity
i) Ratio of depreciation to goods sector backlog
j) Ratio of goods sector backlog to capacity
k) Add goods sector capacity plus capital sector capacity

On a scale of one to ten (with ten being hardest) how hard was this questionnaire?

Answers to Questionnaire (not provided to participants)

Multiple-choice questions:		Matching questions:
1. N	9. A	1. F
2. E	10. C	2. D
3. C	11. E	3. B
4. A	12. Y	4. H
5. C	13. B	5. G
6. D	14. D	6. E
7. D	15. D	7. C
8. B	16. N	8. A

Notes

1. The software for the New interface can be obtained from Louisa Ford (lford-@netcom.ca).

References

Bennett KB, Flach JM. 1992. Graphical displays: implications for divided attention, focused attention, and problem solving. *Human Factors* **34**: 513–533.

Brehmer B. 1992. Dynamic decision making: human control of complex systems. *Acta Psychologica* **81**: 206–223.

Davis S, Meyer C. 1998. *Blur: The Speed of Change in the Connected Economy*. Addison-Wesley: Reading, MA.

Diehl E, Sterman JD. 1995. Effects of feedback complexity on dynamic decision making. *Organizational Behavior and Human Decision Processes* **62**: 198–215.

Dörner D. 1989/1996. *The Logic of Failure: Why Things Go Wrong and What We Can Do to Make Them Right*. Henry Holt: New York.

Fischoff B, Slovic P, Lichtenstein S. 1978. Fault trees: sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance* **4**: 330–344.

Howie E, Sy S. 1996. Interface design of a dynamic decision making simulation (Unpublished BASc thesis). Department of Industrial Engineering, University of Toronto, Toronto.

Paich M, Sterman JD. 1993. Boom, bust, and failures to learn in experimental markets. *Management Science* **39**: 1439–1458.

Richardson GP, Rohrbaugh J. 1990. Decision making in dynamic environments: exploring judgements in a system dynamics model-based game. In *Contemporary Issues in*

Decision Making, Borcherding K, Larichev OI, Messick DM (eds). Elsevier: New York; 463–472.

Sterman JD. 1987. Testing behavioral simulation models by direct experiment. *Management Science* **33**: 1572–1592.

Sterman JD. 1989. Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes* **43**: 301–335.

Sterman JD. 1994. Learning in and about complex systems. *System Dynamics Review* **10**: 291–330.

Sterman JD, Meadows D. 1985. Strategem-2: a microcomputer simulation game of the Kondratiev cycle. *Simulation and Games* **16**: 174–202.

Vicente KJ. 1996. Improving dynamic decision making in complex systems through ecological interface design: a research overview. *System Dynamics Review* **12**: 251–279.

Vicente KJ, Rasmussen J. 1992. Ecological interface design: theoretical foundations. *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-22**: 589–606.

Wickens CD. 1992. *Engineering Psychology and Human Performance* (2nd edn). Harper Collins: New York.

Woods DD. 1991. The cognitive engineering of problem representations. In *Human–Computer Interaction and Complex Systems*, Weir GRS, Alty JL (eds). Academic Press: London; 169–188.