Table of Contents

Go Back

DECISION SUPPORT WITHIN COMPLEX SYSTEMS: AN EXPERIMENTAL APPROACH USING THE STRATEGEM-2 COMPUTER GAME

Lt Col J. ROBERT BOIS, USAF
Air Command and Staff College
Air University, Maxwell AFB, AL
Bob.Bois@knology.net

Abstract

In 1989, John Sterman published his seminal paper, *Misperceptions of Feedback in Dynamic Decision Making*. His misperception of feedback hypothesis deals with the difficulty people have in managing complex environments even when they purportedly have perfect knowledge and have perfect information about the system. Over the years, several authors have attempted to consider how the human failures, which are a prominent part of the misperception of feedback hypothesis, can be reduced. However, these authors have achieved mixed results in attempting to make improvements to human decision support. It is the purpose of the current research to provide meaningful decision support to managers of complex environments. Specifically, the research used the STRATEGEM-2 simulation game and purposely developed a decision support method designed to improve human performance within a complex system. The experiment required subjects to make a single decision within a dynamic system where the task involved feedback delays, nonlinearity of system processes, positive feedback loops, and multiple cues. The decision support included a decision rule and a newly developed game instruction designed to improve participant knowledge and information about the microeconomy of the STRATEGEM-2 simulation. Results of the research have discovered that the new instruction and the decision support rule produced significant results in improving decision making. Additionally, this research demonstrates that the lack of participant motivation levels can mask decision support interventions.

Introduction

Over nearly two decades, John Sterman has researched and written about dynamic decision-making of participants using the STRATEGEM-2 computer simulation game. In a seminal work (1989a), along with several other accompanying articles (1987; 1989b; 1994; 1985), he established that a misperception of feedback in decision environments exists on behalf of participants because they fail to take into account delays between their own decisions and the dynamics of the simulation environment. Further, Sterman suggests that participants are operating with perfect knowledge of the system structure along with perfect information.

Note: This paper is a condensation of the author's dissertation submitted to the University at Albany, State University of New York in June 2002 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy, Information Science Ph.D. Program, Rockefeller College of Public Affairs and Policy. For those who would like the entire study, please feel free to make your request directly with the author at: Bob.Bois@knology.net.

However, given these "perfect" settings, participants consistently perform poorly. Sterman (1989a) developed a misperception of feedback hypothesis misperception of feedback hypothesis from his research.

George Richardson and John Rohrbaugh (1990) essentially challenged Sterman's (1987, 1989a) findings by hypothesizing that if participants were given better cues to consider in the simulation environment they would perform better. They replicated the Sterman (1987, 1989a) study using a changed interface that incorporated revised cue designs. The results, unfortunately were not as predicted. They were mixed – one half of the participants improved their scores while the other half performed worse.

Edward Howie, Sharleen Sy, Louisa Ford, and Kim Vicente (2000), revisited Sterman's (1989a) misperception of feedback hypothesis and once again attempted to improve upon poor participant performance. Howie and others' (2000) approach was very similar to Richardson and Rohrbaugh (1990) with respect to how the simulation information was presented to the participant. Additionally, they expanded their focus to include measuring the level of environment knowledge possessed by each participant. This was an important step forward, in that the assumptions made by "perfect knowledge" had not been tested up to this point. Unfortunately, like the Richardson and Rohrbaugh (1990) experiment, Howie and others (2000) achieved mixed results. However, they did conclude that improving how information is presented to game players does result in improved game scores.

Problem Statement

It is possible that the above findings have not completely resolved the issues surrounding the misperceptions of feedback hypothesis. A major concern is that the misperception of feedback hypothesis puts undue emphasis upon the notion that participants have "perfect knowledge of the system structure along with perfect information." It is more than likely that this cannot be so, and it was demonstrated to some degree by Richardson and Rohrbaugh (1990) and Howie and others (2000). For example, Howie and others (2000) demonstrated that knowledge of the system structure was far from ideal before (and even after) the experiment had taken place. The Howie and others (2000) study, along with Richardson and Rohrbaugh (1990) also made valid criticisms on how the participants of the Sterman (1987, 1989a) studies did not actually have perfect information at their disposal.

Therefore, the problem statement for the current research is: That the explanations of poor performance produced in the Sterman (1987, 1989a) studies may have been flawed, at least to the extent of the "perfect knowledge and perfect information" line of reason. It may be possible, then, to train or aide participants to be better performers.

The problem statement is important for the following reasons: First, Richardson and Rohrbaugh (1990) point out that the information presented in the Sterman (1987, 1989a) studies is less than adequate. Specifically, they determined that in order for participants to make the most out of the information presented by the Sterman (1987, 1989a) simulation, they would require a certain degree of sophistication that most likely would not reside with the average player. In other words, information can be better presented, along with assistance for cue interpretation that should result in improved participant performance.

Second, Howie and others (2000) produce a convincing argument that Sterman (1989a) did not provide appropriate, or adequate, information on the computer display of his simulation.

They suggest that substandard performance on behalf of participants is not due to a lack of knowledge or to psychological limitation.

Third, Howie and others (2000) demonstrated that the premise of "perfect knowledge" does not exist. Participants who were tested a priori and a posteriori exhibited knowledge that was far less than optimal.

Fourth, when preparing for this undertaking, Rohrbaugh♦ suggests that it has become evident that the "setup" of the experiment is equally crucial to the actual experiment itself. Apparently, and too often, researchers provide participants with instructions, and then, "jump" right into the data collection process. In other words, not enough attention has been paid to how participants are instructed. Is it possible then, that participants can be better prepared, or informed, regarding the dynamics of the simulation environment that they are about to take part? Possibly so.

Fifth, and most importantly, it is imperative to learn how to improve dynamic decision-making support. If indeed participants in an experimental setting can learn how to improve their performance with simulated complexity, then it may be possible to design decision support systems to assist real decision makers with the complexities they face in real systems.

Finally, one should not impart from this research that it is an attempt to debunk the misperception of feedback hypothesis. To the contrary, although the misperception of feedback hypothesis may be based, in part, on an incorrect assumption (that participants have perfect knowledge and information), it remains important to realize that human judges have difficulty with delayed feedback systems. Therefore, it is equally important to explore methods that can be used to improve human performance.

In light of the above citations, it is the opinion of this researcher that it is warranted to try "once again" and see if participants can indeed perform better. There are too many "mixed results" requiring further/additional exploration.

Research Hypotheses

Assuming there are ways to improve human performance in the face of time-delayed feedback dynamics, the following hypotheses are projected for this research thesis. They are:

1. If information and knowledge about a system are better understood, participant performance will improve.
2. If participants are provided with a decision rule that focuses their attention on proper cues and how to weigh their importance, their performance will improve.
3. Participants reporting greater effort during the experiment simulation will out-perform those who do not.

Background

In order to commence, the reader must first be made familiar with the mechanics of the STRATEGEM-2 game. The term STRATEGEM stands for a "STRATEgic Game for Educating Managers." STRATEGEMs are a series of games produced for computers. STRATEGEM-2 deals with a micro-economy. It was born from the study of the economic long wave, or Kondratiev Cycle (Kondratiev, 1935). STRATEGEM-2 first appeared in the literature

♦ 2000. Rohrbaugh, J. Personal interview. 4 January.

by Sterman and Meadows (1985) and was developed to teach decision-making dynamics to individual players (or teams) facing positive feedbacks inherent to a Kondratiev Cycle.

Briefly, the game is played as follows: The player is established as a manager for a capital-producing sector of an economy. Game time is divided into two-year intervals beginning with year zero and ending in year seventy. Thirty-five decisions will be required from the player over the seventy-year period. The game board in Figure 1, taken from the original Sterman experiments (1985, 1987, 1989a), is divided into two sectors, a capital sector (in simplicity, this would be the physical/industrial capacity to produce consumer goods and its own capital goods), and a goods sector (you may think of this as the consumer sector). Orders for each sector go into a "backlog of unfilled orders" area where they will sit awaiting shipment to their respective sectors. The amount to be removed from this waiting area is equal to the capacity of the capital stock. Additionally, the capital stock loses ten percent of its level every two years due to depreciation.

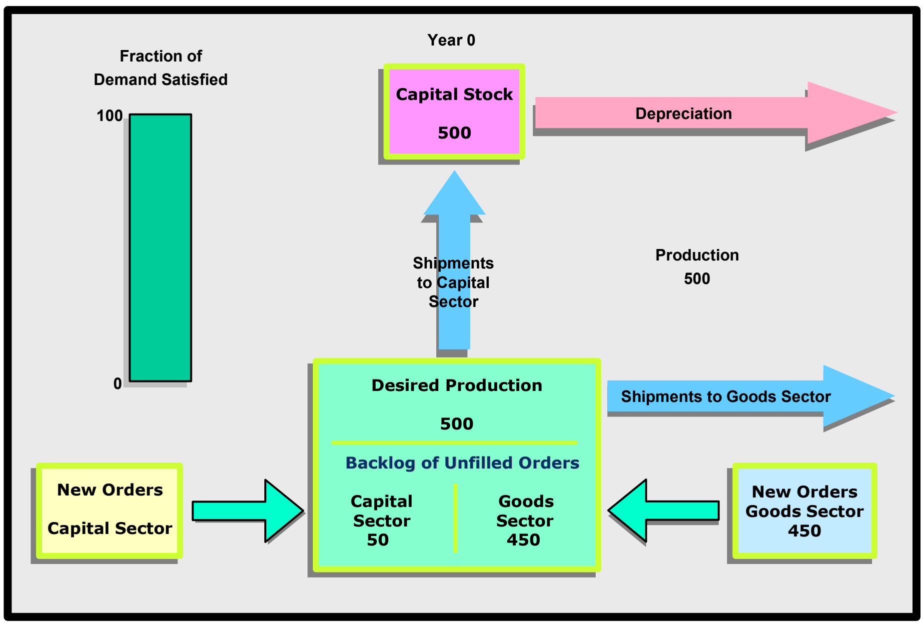


Figure 1 - Sterman STRATEGEM-2 Game Board

In the Sterman experiments (1985, 1987, 1989a), the game begins in equilibrium. This means that the capital stock is at a level of 500. The total for the backlog of unfilled orders is 500 as well (450 unfilled orders for the goods sector and 50 unfilled orders for the capital sector). Finally, a predetermined order of 450 goods sector orders is displayed for the player. This leaves a single decision to be made: How many orders are to be placed in the capital sector? The adept player should be able to recognize that an order of 50 for the capital sector will keep the game in equilibrium. The reason this is so is that 50 units to the capital sector will eventually be used to replace the 50 units of depreciation the capital stock is scheduled to lose (500 ∗ 10%). The combination of these 50 capital sector orders with the established 450 goods sector orders totals 500 units, which is equal to the production capacity of the capital stock. The capital stock will then be able to produce the 450 orders required by the goods sector, and it will be able to produce the 50 orders of capital to replace the 50 units it will lose to depreciation. Therefore, the game will remain in equilibrium.

On the upper left side of the game board is a "thermometer-type" display called, "Fraction of Demand Satisfied" (FDS) -- the bar indicates 100% FDS in year zero. Sterman (1985, 1987, 1989a) also produces a "Production" figure in order to provide information to the player. Production is calculated as the minimum of either capital stock or desired production. Plainly stated, industry would not produce more than demand requires. If the capital stock were larger than desired production, the player would simply be penalized for excess capacity. The FDS bar is merely a function of production divided by desired production. Hence, the only time FDS is less than 100% is when the capital stock is less than the desired production. Figure 2 depicts the STRATEGEM-2 microeconomy from a "stock and flow" perspective used by system dynamists to better show the feedback structure of the economy.

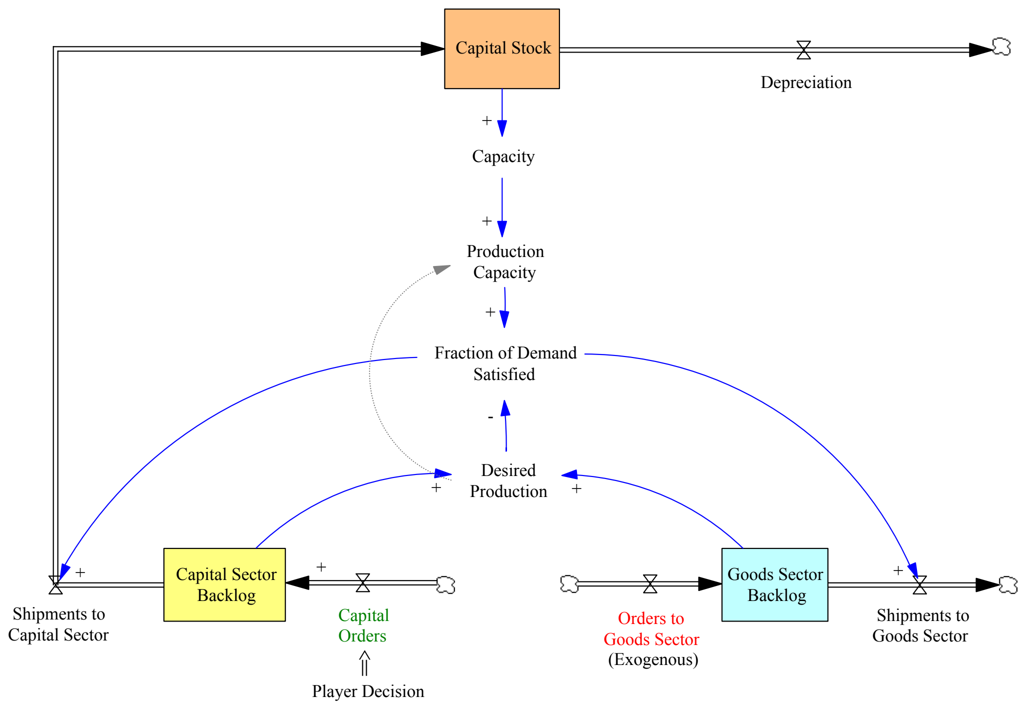


Figure 2 - STRATEGEM-2 Stock and Flow Structure

As a final note on the game mechanics, after each round of play, the game produces a "score" indicating to the player his or her level of performance. The score is a simple mathematical formulation that keeps an accumulating sum of the absolute difference between the total desired production (the total backlog of unfilled orders), and the production capacity of the capital stock (which is equal to the total of the capital stock), and is divided by the time interval (the years of play). For example, after the first round of play, the absolute difference between desired production and production capacity is zero. Divide that by 2 years and the score remains at zero. The score indicates how well each player can balance the interactions of supply and demand. There is equal penalty for excess demand, as well as excess supply.

Complexity Added

To provide complexity to the game, in year four, Sterman (1985, 1987, 1989a) adds a single step increase to orders from the goods sector. Orders go up from 450 to 500 and remain at that level for the remainder of the game (players in the game are unaware of this step increase, or of its longevity♦). The key is that the participant must order more than the depreciation of the capital stock. The reason: The capital stock must be increased in order to meet capacity requirements for the new demand. The "gotcha" of the problem is that the player is more than likely unaware that it will take a few to several years to build up the capital stock to meet the new requirement. Additionally, the increased order to the goods sector further complicates the problem as it continues to grow the backlogs of orders that need to be shipped. This requires

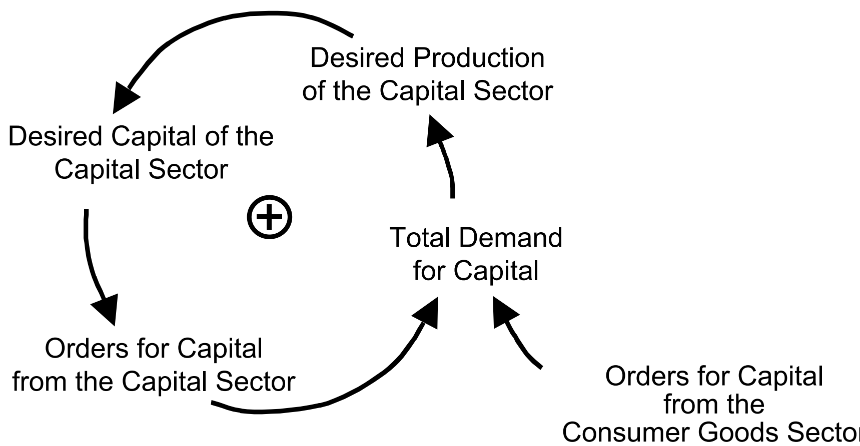


Figure 3 - Multiplier Accelerator Loop
Adapted from (Sterman, 1989a)

that more capital stock be ordered so that the demands of the burgeoning backlog of unfilled orders can be met – a classic multiplier/investment accelerator problem (Figure 3). This positive reinforcing loop in the system normally forces players to order too much capital stock over subsequent years.

Typically, players fail to calculate that with the new increase in orders from the goods sector produces a new equilibrium (the actual new equilibrium raises from 500 to 555 – and will be presented as 560 in the game itself because the simulation rounds to the nearest 10). Therefore, as players build their capital stock to levels well above 560 (trying to counteract the increasing total backlog), they are slow to find that the backlog will quickly drop when capacity to produce is great and the orders for goods sector units remain at 500. When participants in the game realize that they now have too much capital sector inventory, they will tend to stop ordering all together. Depreciation then begins to show its effect by lowering the capital stock. However, the player soon finds him or herself "behind the power curve" once again with not enough capital stock to meet the total demand of the economy -- and the cycle continues. All of these problems are the result of poor anticipation of the delays in the system, as well as not calculating the desired new equilibrium level.

♦ This is a requirement of the experiment. Otherwise, if the subject were to know this information, he or she could possibly plan accordingly and defeat the dynamics the system is trying to simulate.

The Misperception of Feedback Hypothesis

The research performed by Sterman (1989a) has yielded a misperception of feedback hypothesis that attempts to capture why human subjects perform poorly in this simulation game. Simply stated, the misperception of feedback hypothesis occurs in two forms. The first is the misperception of time delays. (Sterman, 1989a, pg. 324). The second form of the misperception of feedback hypothesis comes from decisions to the environment. (Sterman, 1989a, pg. 326). Simply stated, the misperception of feedback hypothesis can be reduced to a subject's failure to appreciate the time delay built into the game, and that they fail to appreciate how their decisions are reflected within the game-playing environment.

Other Pertinent Studies

The Richardson and Rohrbaugh (1990) study, along with Howie and others (2000), had very sound theoretical foundations in challenging a portion of the misperception of feedback hypothesis. Both studies attempt to fill gaps that are perceived to be unexplained in the misperception of feedback hypothesis. Their hypotheses, whether implicit (Richardson and Rohrbaugh, 1990), or explicit (Howie and others, 2000), stated that players in the STRATEGEM-2 game do not have perfect knowledge of their environment, nor does the environment display perfect information. Sterman (1989a) would argue that because the participant can view a graphic screen at anytime during the experiment, they could obtain immediate outcome feedback of what has been occurring in the dynamics of the game. Richardson and Rohrbaugh (1990) provided the exact same outcome feedback as well as provided current depreciation and shortfall information on the game board. Howie and others (2000), out-doing their predecessors, provided all this information on a single game-screen.

However, Howie and others (2000) explicitly brought up a very important facet of the misperception of feedback hypothesis worthy of further consideration – the finding that players, before and after the game, could not demonstrate "perfect knowledge" of the game. Richardson and Rohrbaugh (1990) also grappled with this aspect of Sterman's (1989a) work by asking: "How would players perform if the computer screen directly provided them with the cues appropriate for the task? What effect would different forms of cue presentation have on cognitive learning? These questions are important because they may reveal an alternative explanation for the misperceptions and dysfunctional behaviors found by Sterman (1989a) (Richardson & Rohrbaugh, 1990, pg. 464)."

The fact that Richardson and Rohrbaugh (1990), and Howie and others (2000), arrived at mixed results (however, encouraging), leaves the issues of "perfect knowledge," modern computer interfaces, feedforward cues, and the cognitive learning processes to be unresolved. Therefore, the current study retested the Richardson and Rohrbaugh (1990) precepts. Additionally, the Howie and others (2000), computer interface was used along with the concept of testing participant knowledge.

Another issue that has been brought to the attention of the researcher is by Rohrbaugh. It refers to the "setup" of the experiment to the participants. Too often, according to Rohrbaugh, researchers preoccupy themselves with measuring the many dynamics of their experiments. They bring in human subjects, give them something to read, and then move into the experiment without ever considering whether the setup may have had an impact on the subject's performance. The concept of the setup in the STATEGEM-2 game can possibly have significant feedforward effect. Hsiao (1999) found a study where this procedure was introduced as a distinct form of measurement. "Sengupta and Abdel-Hamid (Sengupta & Abdel-Hamid, 1993) base their research design on the theory of information feedback and provide subjects with three types of computer information feedback: outcome feedback, cognitive feedback, and feedforward. Outcome feedback indicates online numerical reports for important state variables of the software project task. Subjects receiving cognitive feedback have access to online time plots containing the patterns of relevant variables and a tabular summary of these cues. Whereas outcome and cognitive feedback are always available on computer screens, feedforward is conveyed by an hour-long training session prior to the task (Hsiao, 1999, pg. 27)."

Reading the instructions to STRATEGEM-2, used by the three main studies identified in this paper (Appendix A), many questions remained unanswered. Therefore, improvements were attempted to the setup of the experiment that can transfer the dynamics of the game into meaningful knowledge that participants could better grasp and understand.

Research Design

The following research design was used:

1. Used the Sterman instruction in the control conditions presented by Howie and others (2000), (Appendix A)
2. Developed an on-screen tutorial to train game participants (Appendix B)
3. Used the Howie STRATEGEM-2 interface (Appendix C)
4. Tested game knowledge among the participants following train-up
5. Surveyed the participants to determine their level of effort at the end of the experiment (Appendix D)
6. Performed a practice trial, and then two scored trials where orders to the goods sector remained the same (a single step increase in year four) for each trial
7. Enrolled 150 volunteer participants
8. Randomized participants into 4 treatments and conditions as follows:
    a. Receives an on-screen train-up of the Sterman instruction (presented by Howie and others, 2000), a practice trial, a knowledge survey, Q&A, and two measured trials.
    b. Receives an on-screen train-up of the Sterman instruction (presented by Howie and others, 2000), a practice trial, a knowledge survey, Q&A, the Richardson and Rohrbaugh decision rule, and two measured trials
    c. Receives new on-screen tutorial (Bois instruction), a practice trial, a knowledge survey, Q&A, and two measured trials
    d. Receives new on-screen tutorial (Bois instruction), a practice trial, a knowledge survey, Q&A, a practice trial, the Richardson and Rohrbaugh decision rule, and two measured trials

The above treatments and conditions are further explained by the 2 x 2 matrix shown in Figure 4 (below). Along the vertical axis, there are two treatments that received (hypothesized) inadequate/adequate training (the Sterman instruction, or better, not receiving the Bois instruction, and receiving the Bois instruction). The horizontal axis has two treatments that received (hypothesized) non-decision/decision support (no Richardson and Rohrbaugh decision rule and receiving the Richardson and Rohrbaugh decision rule). Four conditions are created from the combinations created by mixing different levels of decision support and game instructions.

Conditions and Treatments	No Rule	Receives R & R Rule
No Bois Instruction	I	II
Receives Bois Instruction	III	IV

Figure 4 - Human Subject Random Group Assignments

Data Collection

The final "score" (described earlier) produced by the simulation game was the main data point captured in this research. The ultimate goal for the participant was to minimize his or her score – the smaller the score, the better the performance.

Additional data points collected included: The scores from the participant knowledge surveys as tested by Howie and others (2000), and the determination of participant level of effort from a self-assessment perspective (Appendix D). Additionally, demographic information was collected and analyzed.

Sample and Subjects

Although this research makes inferences from the sample to the greater population, the researcher used a non-probability/convenience sample of human subjects. Specifically, participants were drawn from graduate/undergraduate students enrolled in the public administration, information science, business administration, finance, and marketing programs at the State University of New York at Albany. In total, 54 graduate and 96 undergraduate students elected to participate in this research. All subjects were recruited on a voluntary basis and did so without receiving any stipend. The participants were randomly placed into the treatments and conditions shown in Figure 4.

Dependent Variables

During the experiment, the dependent variables were the scores received in the first and second trials, the mean average of both trials, (as a reminder, the lower the score, the better the performance for a given trial), and the self-assessed level of effort. The scores indicate the participant's ability to ferret out the important factors in the decision-making process within the dynamic system.

The participant level of effort was surveyed via a self-assessment (Appendix E). This dependent variable has three subsets: self-assessment of individual interest in the research, task understanding, and performance. Following is the assignment of statements to each of the three subset variables:

Variable	**Survey Statement Number**
1. Self-assessment of performance:	3, 7, 8, 12
2. Self-assessment of research interest:	4, 6, 10, 11
3. Self-assessment of task understanding:	1, 2, 5, 9

The questions are designed to get the participants to accurately document their perceptions about their own actions during the experiment. It was presumed that the variables, when analyzed, would reflect on whether they had any significant predictability upon the dependent variables. The survey instrument is based upon a Likert-type scale. It is used to measure the internal states of the subjects (such as attitudes, emotions, and orientations) (Bernard, 2000). The design of each question in the instrument was used to measure whether the participants fully employed themselves during the experiment.

Independent Variables

Independent variables included the game instruction setup, decision support, game knowledge, and demographic information. Operationalization of these variables was as follows: Game instruction setup was derived from subjects receiving either the original Sterman instruction (Appendix A), or the newly devised Bois instruction (Appendix B).

Decision support occurred in two forms. In the first, participants received, or did not receive, the Richardson and Rohrbaugh decision rule. In order to produce this part of the experiment, a card with specific information about the decision rule was provided to those participants destined to receive decision support (Figure 5, below).

----Front Side----

As the manager of the STRATEGEM-2 economy, you have taken it upon yourself to hire a very reputable economic consultant to assist you with your decisions. This person has determined that if you are to follow the formula in the blue box on the reverse side of this card, you will most likely receive an outstanding score for the game. You are reminded by this professional that although you are not required to heed the advice given, you must remain patient and diligent with using the formula.

Example on using the decision aide **in year zero** of the game:

1. Take the current depreciation of 50 units and multiply it times 2 (for 100).
2. Add to that the shortfall* (currently 0) and divide by 2 (which equals 0).
3. Then subtract the current capital backlog (*not total backlog*) of 50.
4. This produces an order of 50 capital units for Year 0

*** Shortfall** = (total backlogs - current capacity).
If this figure computes to less than zero, use zero.

----Back Side----

--- To Order ---

1. Plan in advance to replace depreciation loss

(DEPRECIATION x 2) = _______

2. Shortfall: Reconcile total backlogs with current capacity

add (SHORTFALL ÷ 2) = +_______

3. Adjust for prior orders not yet filled

subtract (CAPITAL BACKLOG) = - _______

Total Orders = _______

Figure 5 continued - Richardson and Rohrbaugh Decision Rule Input Card

In the second form of decision support, participants received, or did not receive, the Bois instruction (Appendix B). This instruction was designed as an on-screen tutorial and had two learning inducement objectives in mind during development, 1) to get participants to understand how STRATEGEM-2 is played, the different features of the game board, and what information is being conveyed to the participant from the various features of the interface, and 2) to get subjects to understand the concept of "equilibrium" within the game dynamics. The tutorial was set up as a linear program to introduce the various teaching elements and included a branching design as each successive page of the tutorial unfolded for the participant. Additionally, criterion frames were used to examine/test participant knowledge of the equilibrium concept and provided direct feedback in order to assist in the learning process. As a final note, close attention was paid to the passage lengths of each of the tutorial's pages so not to overtax subject attention spans.

Game knowledge was tested by adapting the Howie and others (2000) knowledge survey. Scoring of this survey was based upon the number of correct responses on a 0 to 100 percentage scale. Regarding demographics, operationalization of this variable included participant: gender, age, graduate status, years of professional experience, total time on task, and test scores (from the knowledge survey).

Experiment Setup

For all conditions surveyed, the setup of the experiment included either the original Sterman instruction (Appendix A) or the Bois instruction (Appendix B) along with an overview of the Howie STRATEGEM-2 interface (Figure 6). Additionally, the conditions either included, or did not include, the Richardson and Rohrbaugh decision rule. After a train-up session was conducted, a practice trial of the game was played by each subject followed by the knowledge survey and then a question and answer session.

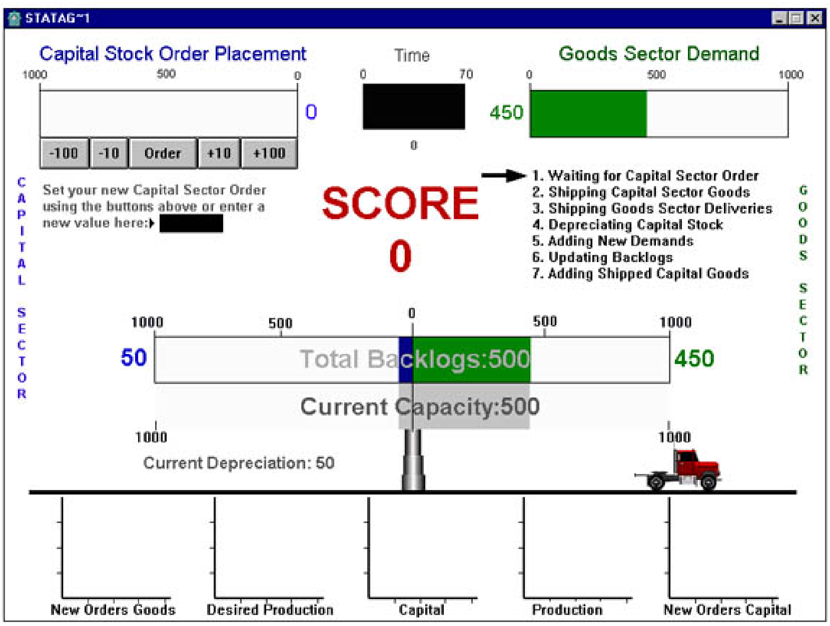


Figure 6 - Howie STRATEGEM-2 Interface

Data Analysis

The data analyses included simple descriptive statistics that were used to capture the broad spectrum of data points among the participants. The main analysis performed was a 3-way analysis of variance. It was used for comparison of the instruction set (receives the Sterman instruction – no Bois instruction – or receives the Bois instruction) put against the decision support rule (receives or does not receive the Richardson and Rohrbaugh decision rule), and further compared with a measure of self-reported motivation. The analysis was used to determine the main effect of the Bois instruction, the Richardson and Rohrbaugh decision rule, and motivation level upon participant performance.

Data Reduction

When considering the entire data set after all collection had been completed, it became obvious that some scores obtained in the two recorded trials were so high, that some form of reduction, or elimination of cases, would be required in order to better capture the true performance of the body of participants, and attempt to reduce or eliminate problems associated with regression to the mean. For example, over two thirds of all participants scored less than 1,000 points for either Trial 1 or Trial 2 (Reminder: The lower the score the better the performance. The Sterman optimal score for the game is 19, and the Richardson and Rohrbaugh decision rule produces and optimal score of 67). Additionally, three subjects scored in excess of 10,000 points in both Trials 1 and 2.

The researcher has determined that individuals receiving very high scores possibly did not understand the task, or they failed to grasp the requirements of the instructions. In order to set some sort of demarcation, any case with a Trial 1 or Trial 2 outlier in excess of 4,000 points was eliminated from the data set. Therefore, 12 cases were eliminated from the original 150, reducing the total N to 138, or by 8 percent.

Data Conversion

In order to better visualize the score data obtained during the experiment, Figure 7 shows boxplots for Trial 1 (T1), Trial 2 (T2), and the two-trial average (TA) for the four conditions generated from the various treatments. Additionally, it demonstrates the existence of several mild and extreme outliers of the raw scores.

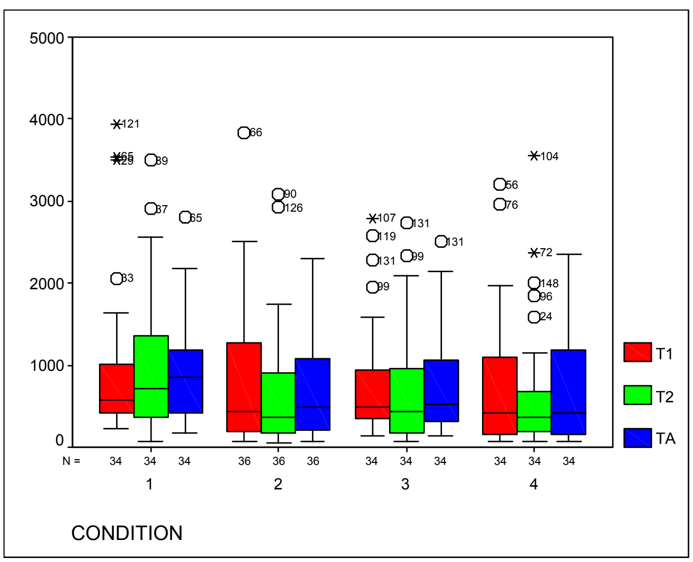


Figure 7 - Raw Score Boxplots by Group

As can be seen in this view for data depiction, the scores produced for the participants for Trial 1, Trial 2, and the two-trial average had large ranges, coupled with their large standard deviations (some even larger than their means), a method to compress the data was searched for that could effectively convert the data in hopes of reducing the large size of the standard deviations and reducing the number of outliers. Therefore, transformations of the data that were attempted included square/cube root conversions and logarithmic conversions. Using a Base10 Logarithmic conversion of the scores proved to provide the best compression of the data and elimination of outliers, while at the same time, maintaining the integrity of how the data relates to each other among the various treatment groups. Figure 8 shows the compressed data. Again, the lower the Base10 Logarithmic score, the better the performance.

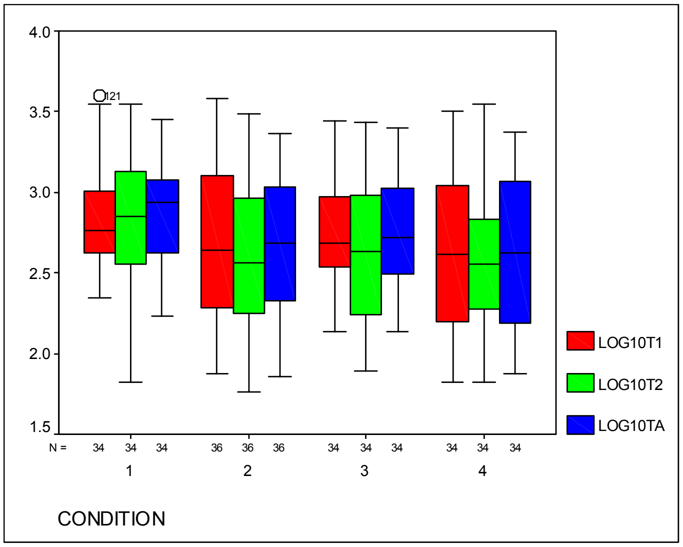


Figure 8 - Base10 Logarithmic Transformation of Scores by Group

Data Description

Descriptive information for all variables (a total of 54 data points were collected for each participant) is not shown in this section. Only pertinent variables that may have some bearing on the research are discussed. Descriptive information for pertinent variables, along with their coding, is as follows:

1. Gender: Male = 1, Female = 2
2. Status: 1 = Undergraduate student, 2 = Graduate student
3. Log10T1[♦]: Base10 logarithmic conversion of the T1 score
4. Log10T2[♦]: Base10 logarithmic conversion of the T2 score
5. Log10TA[♦]: Base10 logarithmic conversion of the TA score
6. TS: Test score (knowledge survey result)
7. SA3: Self-assessment survey question 3 (1 to 5 scale – 1 represents strongly disagrees, 5 represents strongly agrees)
8. SA3FIVE[††]: Identifies participants that scored SA3 with a "5"

[♦] These variables were not "collected," rather, they were computed within SPSS.

Motivation Factor

At this point, special emphasis needs to be made regarding how the level of effort was operationalized during the data analysis process. Hsiao (1999) discovered only three methods of measuring "level of effort" on behalf of participants in a dynamic decision-making (DDM) study. They are: First, is the amount of decision time (how long does it take to make a decision). Second, is the amount of information use for specific information items (is the participant using the information provided in the experiment). Third, is the amount of discussion among participants (do they seek each other's help when allowed by the experiment).

This researcher, interested in this aspect of DDM, posits that if human subjects really tried hard, they would perform well with respect to the various treatments they are exposed to in the current experiment. The idea was to administer a post-experiment self-assessment survey where subjects would be able to self-identify: 1) how hard they were trying, 2) their knowledge of the game, and 3) their interest in the research project.

After several analyses, it was discovered from the subjects' self-assessment survey that their "task knowledge" and/or their "interest in the research" were not good predictors of their effort. However, the statements regarding their performance in the self-assessment survey may have been somewhat ambiguous – except for one statement. The variable, SA3 (self-assessment survey item #3), stated: "I did my best in performing during this experiment;" the position of this statement establishes that if someone was really trying hard, he or she would give this a top rating of "5" (meaning that they strongly agree with the statement). It is believed that this one measure alone can identify a subject who was "motivated." All others ranking this statement less than "5" is considered to be unmotivated, or at least, not as motivated as the researcher would like them to be.

Extending the logic of motivated vs. unmotivated, the boxplots in Figure 9 show a marked difference from the boxplots shown earlier for all cases. Here, the motivated individuals by group have been separated from those who are unmotivated. Clearly, from a descriptive point of view, the differences in performance between those who are motivated and those who are not appear to be noteworthy, and warrant further investigation.

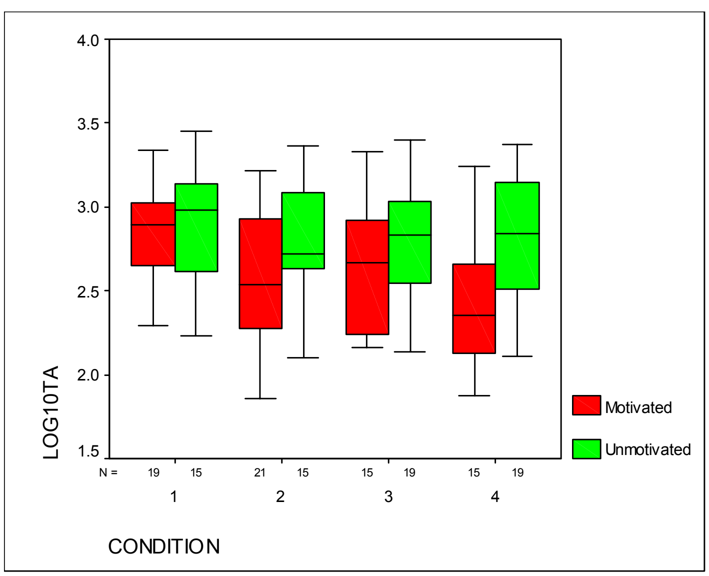


Figure 9 - Performance Comparisons s of Motivateds vs. with Unmotivateds

Findings: Descriptives

Considering the descriptive statistics for all participants (Table 1), the gender difference is near evenly split (53% female). Graduate students made up 34% of participants tested. Test score data (results of the knowledge survey) ranged from 26 to 91 with a mean of 55 and was evenly distributed. Regarding the Base10 Logarithmic scores obtained, the three variables measured have standard deviations that are very small compared to their means. The final two variables, SA3 and SA3Five simply show the range of motivation (SA3) and that the number of motivated participants (SA3Five) represented 51% of the sample population.

Descriptive Statistics of Pertinent Variables

	Minimum	Maximum	Mean	Std. Deviation
Gender	1	2	1.53	.50
Graduate Level	1	2	1.34	0.48
Log10T	1.83	3.60	2.73	.40
Log10T2	1.76	3.55	2.66	0.45
Log10TA	1.86	3.45	2.74	0.39
SA3Five	2	5	4.38	0.74
SA3Five	0	1	0.51	.50

N = 138

Table 1 - Descriptives for All Participants

It was important, however, to determine if the participants were randomly distributed among the four conditions established by the method of study (Bois instruction vs. no Bois instruction, and Richardson and Rohrbaugh decision rule vs. no decision rule). To do so, a one-way analysis of variance was conducted for each of the variables in Table 1 (cross-checked against each of the four research conditions) to see if any non-random assignments could be found as significant ($p < .05$). The result of this test indicated that no variable was found to have a significant non-random assignment. This means that the assignment of participants to the various treatments was indeed statistically random.

Analysis of Variance

The first research hypothesis, improving knowledge and information, is represented by applying the Bois instruction. The second hypothesis, focusing participant attention on proper decision cues and weights, is identified by the application of the Richardson and Rohrbaugh decision rule. The last hypothesis, participants reporting a greater level of effort, is not directly reflected in Figure 4, however, it was included as a third factor in the analysis of variance.

The main effects observed in an analysis of variance are shown in Figures 10 through 12 (below). Three ANOVAs were performed. They included analyses of the Base10 Logarithmic scores of the first trial, second trial, the two-trial average. The predominant trend that is seen in the following figures is that the mean performance scores for the first and second trials, along with the two-trial average, show improvement when either the Bois instruction, or Richardson and Rohrbaugh decision rule, is applied. Additionally, there is a pronounced improvement in scores on behalf of participants who were assessed as motivated over those who were not.

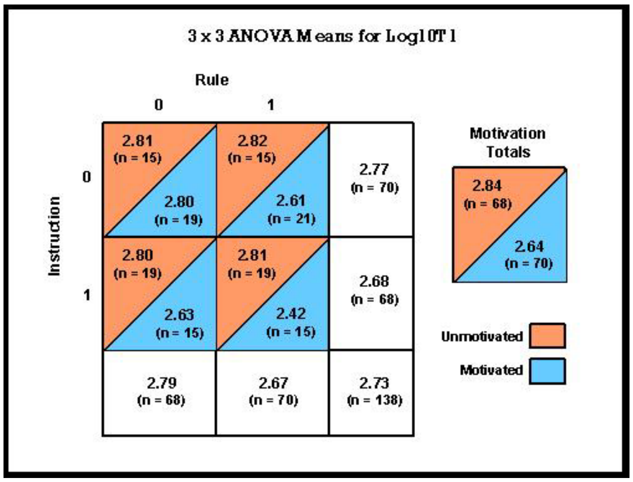


Figure 10 - ANOVA Findings for Trial 1

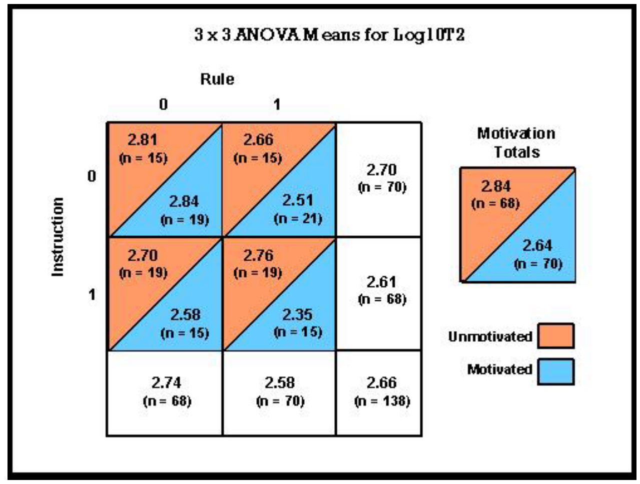


Figure 11 – ANOVA Findings for Trial 2

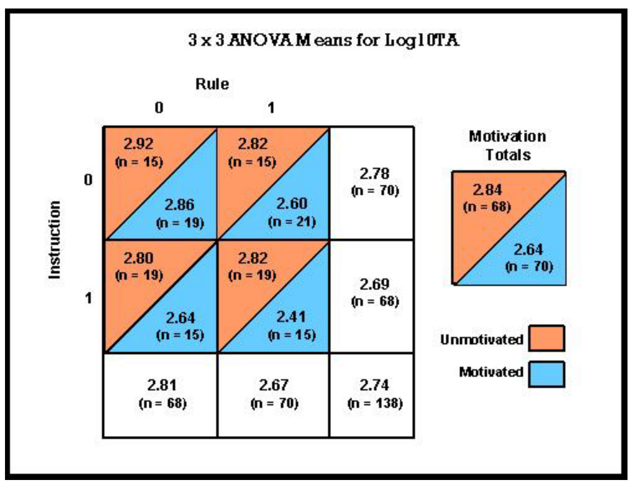


Figure 12 - ANOVA Findings for the Two-Trial Average

In an attempt to better show (graphically) the results of the three ANOVAs, the following Figures 13 through 15, demonstrate the results of each of the three analyses. What is important to remember is that the circles represent participants not receiving the Bois instruction, the triangles represent the reception of the Bois instruction, and the left aligned circles and triangles represent participants not receiving the Richardson and Rohrbaugh decision rule (compared to those aligned on the right side of the chart – they received the rule). Motivation is also separated by color as indicated.

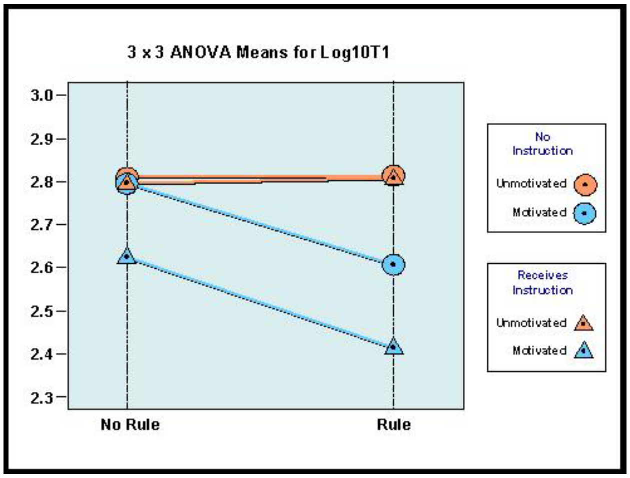


Figure 13 –Graph for Trial 1 ANOVA

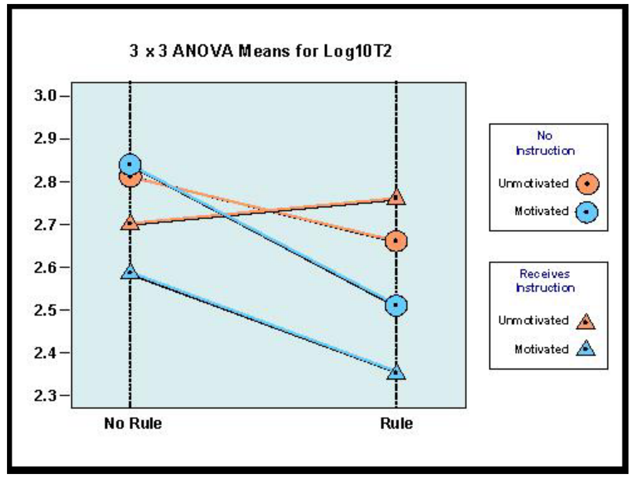


Figure 14 –Graph for Trial 2 ANOVA

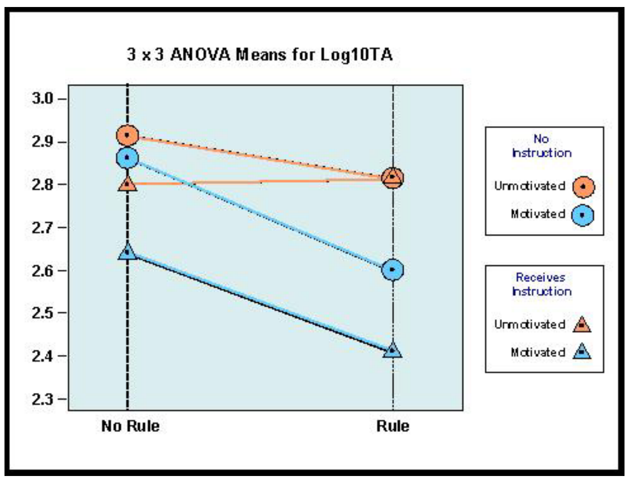


Figure 15 - Graph for Two-Trial Average ANOVAe

What is important to discern in Figures 13 through 15 is the difference in performance between the motivated/unmotivated subjects. For example, the unmotivated subjects show little or no improvement between those who received the Bois instruction and those who did not. The same relationships can be discerned between those subjects receiving the Richardson and Rohrbaugh decision rule to those who did not. However, what is critically important to observe, is that among motivated subjects, the differences in performance between those who received the Bois instruction and those who did not, along with the comparison of subjects receiving, or not receiving, the Richardson and Rohrbaugh decision rule, all perform as hypothetically predicted.

Table 2 (below) shows the F-ratios obtained in the analysis of variance. All three main effects were significant when the two-trial average was used as the dependent variable. Motivation also had a significant main effect in the first and second trails. Additionally, the rule had a significant main effect in the second trial.

F-Ratio of Main Effects

	Log10T1	Log10T2	Log10TA
Instruction	3.11	2.11	4.13*
Rule	3.31	4.72*	5.05*
Motivation	11.07**	4.74*	10.93**

* Sig. at the .05 level
** Sig. at the .001 level

Table 2 - F-Ratios of Main Effects for Instruction, Rule, and Motivation

Motivation Factor Explained

In the third hypothesis, "participants reporting greater effort will out-perform those who do not," two two-way analyses of variance were performed to determine the significance of the Bois instruction and the Richardson and Rohrbaugh decision rule with those who are motivated, and those who are not. Table 3 shows the results of these two analyses.

Log10TA Motivation F-Ratios

	Unmotivated	Motivated
Instruction	.44	4.80*
Rule	.25	7.02**

* Sig. at the .05 level
** Sig. at the .01 level

Table 3 - Motivation F-Ratios

As a final addendum to this section, another very interesting discovery was made when comparing participant knowledge survey scores to their self-assessed motivation levels. The boxplots in Figure 16 (below) show a very different level of performance in test scores (knowledge survey) between the two motivation levels. Additionally, in Table 4 (below) are the descriptives for these two levels of measurement, along with a two-tailed significance test of their means. The means differences between them are not only large (10 points), but their significance is at the .0001 level. The significant differences in test scores may attribute, in some way, to the increased performance of simulation scores between the two motivation levels of subjects.

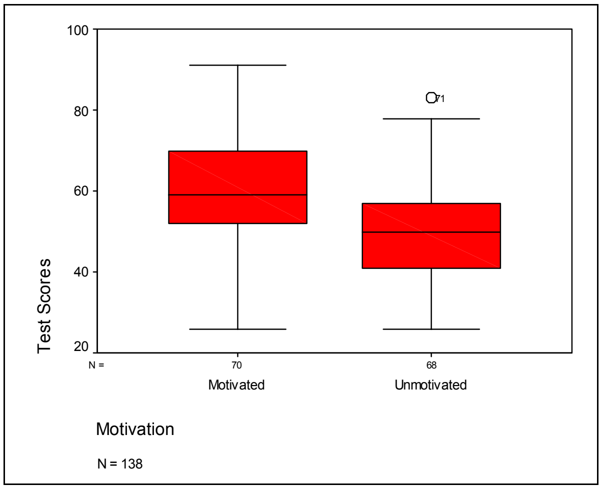


Figure 16 - Test Scores by Motivation

Motivation Comparisions

	Mean	Std. Deviation
Motivated	59.61	15.00
Unmotivated	49.75	13.55
Mean Difference	9.86****	

****Sig. at the .0001 level

Table 4 - Two-Tailed T-Test of Motivation between Groups

Anecdotal Observations

Given the statistical findings of this research, other observations about the experiment must be highlighted. For example, although it is not quantified, participants who received the Richardson and Rohrbaugh decision rule used it in different ways. The researcher observed that when given the rule card, participants at times would simply discard it. At other times, they would try to perform the calculations prescribed by the card, only to abandon the rule card over time. Yet, others would follow the prescriptions of the rule to the very end of the experiment.

The "score" in the simulation was also another area of concern. It seemed that several participants would focus too much attention on this output of the game interface. For example, several participants would preoccupy themselves with trying to obtain a lower score versus trying to properly balance supply and demand.

Depreciation did not seem to be fully understood by many participants. It is the ONLY means of reducing capital stock. In other words, when current capacity was too large for the desired production, there were several subjects who neglected to simply order "zero," in order to lower their capital stocks. Many participants failed to appreciate that during times of excess capacity, depreciation could be used to assist them in lowering their production capabilities in order to try and balance their supply with demand.

Finally, as an overall observation, many participants found the simulation very difficult to understand. This is from an observational point of view and could not be corroborated with self-assessment data.

Ambiguities

When considering the questions posed during the self-assessment survey portion of the experiment, it was discovered that those statements dealing with "task knowledge" and "research interest" where not of any statistical value. However, statements dealing with self-assessment of "performance," several ambiguities were discovered that may have led participants to misunderstanding what exactly was being presented. For example, self-assessment statement #7 says, "When provided with a set of decision cues to follow, I followed them all the time." The problem with this statement is the word "cues." What does it mean? Is it likely that the average participant would not understand what is being stated? Additionally, this statement was geared toward subjects who had received the Richardson and Rohrbaugh decision rule, and/or subjects who received the Bois instruction. These participants were pointed to specific elements of the simulation and how to react to them; however, they were never told that these elements were "cues." This can lead to very inappropriate understanding of the statement. Self-assessment statements #8 and #12 were found to have similar ambiguities. Only statement #3 was discerned to be unambiguous.

The only other item that can be considered ambiguous deals with the verbiage of the Richardson and Rohrbaugh decision rule card that was used for participants receiving the rule. The card (see Figure 6 above) has two sides. On the first side, the participant is told that they have hired a reputable consultant to assist in STRATEGEM-2 decision making. The participant is reminded that they must remain diligent with using the decision formula presented by the consultant (which is the Richardson and Rohrbaugh decision rule). A sample "work through" of the rule is also presented on the front side of the card. On the reverse side of the card is a layout of the formula that the participant can use by simply "plugging in" numbers found on the game interface. The layout then provides a step-by-step process whereby the participant then arrives at a calculated game input – a number to be used for capital goods orders.

Several ambiguities were discovered after the fact that has led the researcher to wonder how effective the treatment of the Richardson and Rohrbaugh decision rule was upon game play. For example, on the front and reverse side of the card, the term "shortfall" was clarified for the participant. Directly below this statement was added verbiage stating: "*If this figure computes to less than zero, use zero.*" An ambiguity occurs because this added statement was meant to relate to the computed final "total orders" produced by the Richardson and Rohrbaugh decision rule and not to the "shortfall" amount. Additionally, more ambiguity occurs because the rule does not address when computations end with a value that is not evenly divisible by 10 (because the game interface rounds all values to their nearest 10).

The final ambiguity discovered was on the reverse side of the card whereby the shortfall amount was shown to be added to the computed depreciation value. This is correct; however, if the shortfall computes to a negative number, the participant needs to know that instead of adding, they would now be subtracting the shortfall amount from the computed depreciation value.

Given these findings regarding ambiguity with the Richardson and Rohrbaugh decision rule, the researcher was uncertain as to what their effects are on the results of those treatment groups that were exposed to the rule. The reason being is that in the face of the ambiguities, several participants were able to use the rule card and achieved very low scores. Others did not, but was that a result of the ambiguities, or that possibly they simply discarded the rule (as was observed by the researcher as an anecdotal finding), or was it that they were simply not analytically inclined to fathom the directions proposed by the Richardson and Rohrbaugh (1990) formula? These questions cannot be fully resolved. However, as a minimum, mean scores of the treatment groups using the Richardson and Rohrbaugh decision rule were at a level consistent with hypothetical predictions (regardless of their statistical significance). Findings for these data, therefore, will remain as stated.

Conclusions

First Hypothesis: The Impact of Knowledge and Information

The first hypothesis in the research postulated: If information and knowledge about a system are better understood, participant performance will improve. The control for this hypothesis was represented by participants not receiving the Bois instruction. The treatment was to introduce the Bois instruction to another set of randomly assigned subjects. The research question associated with this hypothesis asks: Can proper/adequate knowledge and information about the system be taught to participants? The intervention of a new instruction, one that teaches participants on the necessary knowledge to become a better decision maker within the dynamic environment, has shown to have had a significant effect ($p = .05$) towards improving decision-maker performance. This was observed from the main effect that the instruction had upon the two-trial average score. The significant F-ratios found for the mean average two-trial performance suggest that this may be so. However, caution must be exercised. For example, were these improved performance scores due to iteration? Cognitive style? Or, participant learning style? The answers to these questions are not known from the current study as these areas of interest were not measured during the experiment.

Second Hypothesis: The Impact of Decision Support

The second hypothesis of the study states: If participants were provided with a decision rule that focuses their attention on proper cues and how to weigh their importance, their performance will improve. The control for this hypothesis was represented by participants not receiving the Richardson and Rohrbaugh decision rule. The treatment was to introduce the Richardson and Rohrbaugh decision rule to another set of randomly assigned subjects. The research question associated with this hypothesis asks: Can participant performance be improved via decision cues and weights? The intervention of a decision support rule, one that directs participants toward specific cues and provides a weight for their importance, has shown to have significant effects ($p = .05$) toward improving decision maker performance. The rule's main effect was significant for the second trial score, as well as the two-trial average. As in the first hypothesis, the significant F-ratio scores for the two-trial average suggest that improvements to the decision-making process can be made through the use of cues and weights. Again, did iteration, cognitive style, or learning style have a factor in this finding? The answer to this question cannot be determined from the current research design.

Third Hypothesis: The Impact of Level of Effort

The final hypothesis of the study suggested: Participants reporting greater effort during the experiment simulation will out-perform those who do not. This hypothesis is used in an attempt to answer the following research question: Can a participant's self-assessment of level of effort be used to better determine their own experiment performance? Level of effort was operationalized through motivational self-assessment on behalf of the participant. The discoveries made in this area were found to have a noteworthy impact upon experiment results.

Participant self-assessed motivation had a significant effect ($p = .05$) on the second trial score, and it had very significant effect ($p = .001$) on the first trial and two-trial average scores. The motivation factor was also found to have a masking effect upon the results. This means that by measuring the subjects' motivation level, a truer overall picture of performance can be obtained. This was shown when a performance comparison was made between the motivated and unmotivated subjects. The first of these discoveries was found in the significant F-ratios for each trial (and the two-trial average) of the experiment. These ratios suggest that subjects who really try hard to implement experiment interventions consistently have a greater effect (improved performance) than those who do not. This is an important finding in light of the corpus of the dynamic decision-making literature for it posits the question of how important other research findings have been because they have not been filtered/differentiated for motivation factors.

The angle of this specific portion of the research is to determine if there is a masking of the results obtained that can somehow be peeled away, revealing a better understanding of the experiment treatments. Specifically, when the data set was divided between motivated and unmotivated participants (as self-assessed from the viewpoint that "they did their best" while participating in the experiment), it was found that those who self-assessed themselves to be motivated – having done their best – outperformed those lacking full motivation.

For the motivated subjects, significant F-ratio results were found for the motivated participants versus the unmotivated participants. Therefore, it is possible that lower motivation levels (those that are not fully motivated) mask the intended treatments that are designed to improve decision making in the STRATEGEM-2 environment.

Using the motivation discriminator reveals another interesting facet of the research. Test scores (results of the knowledge survey) averaged about 55 percent (on a 0 to 100 percentage scale) for all participants. Yet, when considered independently between those subjects that were motivated and those that were not, the mean scores were about 60 percent for motivateds versus 50 percent for unmotivateds. This was a clear indication that the motivational level produces improved results upon performance, and it was found significant at the .0001 level.

Discussion

This research project began with the notion that the Sterman (1989a) experiment with STRATEGEM-2 may have been flawed with respect to the misperception of feedback hypothesis. Specifically, participants in the simulation performed poorly in light of having "perfect knowledge and perfect information" while undergoing the rigors of play.

It is the current research initiative that the Sterman (1989a) observations regarding the misperception of feedback hypothesis remain accurate to some degree. This means that participants perform poorly because they often do fail to properly perceive the time delays in the system, and they fail to understand the effect of their decisions to their environment. These elements of the misperception of feedback hypothesis cannot be eliminated from current findings, however, what cannot be corroborated, is the perfect knowledge and information premise made by Sterman (1989a). For example, as was performed in the Howie and others (2000), knowledge of the simulation and system environment was tested in the current study. The results in this portion of the experiment once again clearly demonstrate that participants do not possess perfect knowledge of the system. Regarding whether participants possess perfect information is also debatable. Although Howie and others (2000) were able to demonstrate how an improved simulation interface works toward improving the information about the system to the participant and, in turn, contributes toward better participant performance, one cannot say that the information presented is perfect. This facet was not tested by Howie and others (2000), or by Sterman (1989a), yet was claimed to exist by Sterman (1989a). The current study does not profess that such an ideal of "perfect information" exists, and it cannot be determined how such a concept can even be measured.

The current research argues that the notion of perfect knowledge and information should no longer be a part of the misperception of feedback hypothesis. Rather, the opposite is more probable, that perfect knowledge and information are not a benefit enjoyed by participants.

Given that test subjects do not have perfect knowledge and information, it remains to know if they can be *taught* to make better decisions (the Bois instruction), or can they be *shown* to make better decisions (the Richardson and Rohrbaugh, decision rule). It is felt that this occurred on both counts – particularly when participants where screened for self-assessed motivation levels. However, caution is warranted; for it is unknown if the improvements observed from the interventions were not a result of other issues that were not measured (e.g. cognition, learning, and iteration). Given that significant effects in decision making were recorded for all two-trial average (Log10TA) scenarios, the results are still encouraging that either the Bois instruction, or Richardson and Rohrbaugh decision rule, were able to assist decision makers improve their performance over those subjects that lacked any assistance at all. The misperception of feedback hypothesis remains an important barrier towards effective decision making in dynamic environments; however, this study shows promise that decision makers can be aided in improving their decision-making skill in these environments.

The findings from the current research indicate three important factors that can be used to improve decision-making support in dynamic environments. The first factor is motivation. Clearly, this factor produced significant results across all levels of the simulation and it is important to take notice of it. Decision support researchers and consultants need to begin paying attention to this factor. Because the lack of motivation has a tendency to mask intended decision support interventions, it is imperative that decision support systems, particularly those in real world environments, consider ways to motivate decision makers to become motivated at a very high level. The methods to do so are undetermined from the perspective of the current research. However, they could include such things as: monetary reward, enlisting decision makers to have a greater "stake in the outcome," and improved benefits (health coverage, retirement benefits, insurance coverage, compensation time-off, improved workspace, to name a few). This list is not exhaustive, yet, provides consideration for improving motivation among real-world decision makers who are operating in dynamic environments. For researchers, it represents a possible list of factors that can be used to determine the effectiveness of improving decision-maker motivation.

The second factor that can be used to improve decision-maker performance is instruction. The current research focused on increasing participant knowledge and interpreting information within a simulation environment. It is posited that the same can be translated to a real-world environment. Researchers and consultants in this area would need to focus more attention on trying to explain the dynamics of decision environments to decision makers. For example, in the current research, the Bois instruction focused very heavily on explaining the concept of "equilibrium" in the STRATEGEM-2 environment. This concept is not unique to STRATEGEM-2, but is applicable to most dynamic decision environments.

The third factor that can be used to improve decision-maker performance is rule based. The Richardson and Rohrbaugh decision rule not only had a significant effect on experiment results, it is possible that it provided great benefit to decision makers who had most difficulty in understanding the STRATEGEM-2 environment. It is opined that, possibly, decision makers who are most inclined to approach decision situations in an intuitive (cognitive) manner would benefit most from such a decision rule. This is opposed to an analytically inclined decision maker who would rely more on his, or her, understanding of the environment to make better decisions. The effect of the Richardson and Rohrbaugh decision rule cannot be truly appreciated from the current research because cognitive faculties were not measured on behalf of the participants. However, a measurement of cognition among decision makers may possibly improve decision support interventions. This area is imperative for future research and should not be overlooked by real-world decision-support consultants.

References

Bernard, H. R. (2000). *Social Research Methods: Qualitative and Quantitative Approaches*. Thousand Oaks, California: Sage Publications.

Howie, E., Sy, S., Ford, L., & Vicente, K. J. (2000). Human-computer interface design can reduce misperceptions of feedback. *System Dynamics Review, 16*(3), 151-171.

Hsiao, N. (1999). *In search of theories of dynamic decision making: A literature review.* Paper presented at the International System Dynamics Conference, Wellington, New Zealand.

Kondratiev, N. (1935). The long waves of economic life. *Review of Economic Statistics, 17*, 105-115.

Richardson, G. P., & Rohrbaugh, J. (1990). Decision Making in Dynamic Environments: Exploring Judgments in a System Dynamics Model-Based Game. In K. Borcherding, O. I. Larichev, & D. M. Messick (Eds.), *Contemporary Issues in Decision Making* (pp. 463-472). Amsterdam: North-Holland.

Sengupta, K., & Abdel-Hamid, T. K. (1993). Alternative Conceptions of Feedback in Dynamic Decision Environments: An Experimental Investigation. *Management Science, 39*(4), 411-428.

Sterman, J. D. (1987). Testing behavioral simulation models by direct experiment. *Management Science, 33*, 1572-1592.

Sterman, J. D. (1989a). Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes, 43*(3), 301-335.

Sterman, J. D. (1989b). Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science, 35*(3), 321-339.

Sterman, J. D. (1994). Learning in and about complex systems. *System Dynamics Review, 10*(2-3), 291-330.

Sterman, J. D., & Meadows, D. (1985). STRATEGEM-2: A microcomputer simulation game of the Kondratiev Cycle. *Simulation and Games, 16*(2), 174-202.

APPENDIX A

The Sterman Instruction

Welcome to the STRATEGEM-2 Simulation Game♦

Version 2.1 Copyright 1985 John Sterman

The economic malaise of the 1980's has revived interest in the economic long wave or Kondratiev Cycle, a cycle of prosperity and depression averaging 50 years.

Since 1975 the System Dynamics National Model has provided an increasingly rich theory of the long wave. The theory emerging from the National Model explains the long wave as the endogenous result of decision making by individuals, corporations, and government. However, the complexity of the National Model makes it difficult to explain the dynamics underlying the long wave. This game demonstrates how long waves can arise by focusing on the role of capital investment.

There are two basic kinds of industries in modern economies: capital producers and producers of consumer goods and services. Goods producers sell primarily to the public. Producers of capital make and sell the plant and equipment that the consumer sector needs in order to produce goods and services. But, in addition, the capital-producing industries of the economy (construction, heavy equipment, steel, mining, and other basic industries) supply each other with the capital, plant, equipment, and materials each need to operate. Viewed as a whole, the capital sector of the economy orders and acquires capital from itself.

You will manage the capital producing sector of the economy. Your goal is to balance the supply and demand for the capital. To do this you must keep your production capacity (current capacity) as closely matched to the demand (total backlogs) for capital as possible. The game is won by the person with the lowest score. The score is the average absolute deviation between production capacity and desired production. For example, if capacity were 500 and demand were 600, your score for that period would be 100. Likewise, if capacity were 600 and demand were only 500, your score for that period would also be 100. A Score of zero means supply and demand are in perfect balance. You are therefore penalized for excess capacity (which implies some of your factories are idle) and also for insufficient capacity (which means you are unable to meet the demand for capital).

Time is divided into two-year periods. At the beginning of each period, orders for capital are received from two sources: the goods sector and the capital sector itself.

Orders for capital arriving from the goods sectors are determined by the computer. Orders for capital you placed in the previous period are moved into the unfilled order backlog for the capital sector.

Orders placed by the goods and capital sectors accumulate in the backlog of unfilled orders for each sector. The total backlog of orders is the desired production for the current two-year period, the demand you must meet.

Production itself is the lesser of desired production or production capacity. Production capacity is determined by the capital stock of the sector. Capital stock is decreased by depreciation and increased by shipments. You lose 10% of your stock each period.

If capacity is inadequate to meet demand fully, available production of capital is allocated between the capital and goods sectors in proportion of their respective backlogs. For example, if the backlog from the capital sector were 500 and the backlog from the goods sector were 1000, desired production would be 1500.

If capacity were only 1200, production would be 1200 and the fraction of demand satisfied would be 1200/1500 = 80%. Thus 400 units would be shipped to the capital sector and 800 would be shipped to the goods sector.

Any unfilled orders remain in their respective backlogs to be filled in future periods. In the example, 100 units would remain in the backlog of the capital sector and 200 would remain in the backlog of the goods sector.

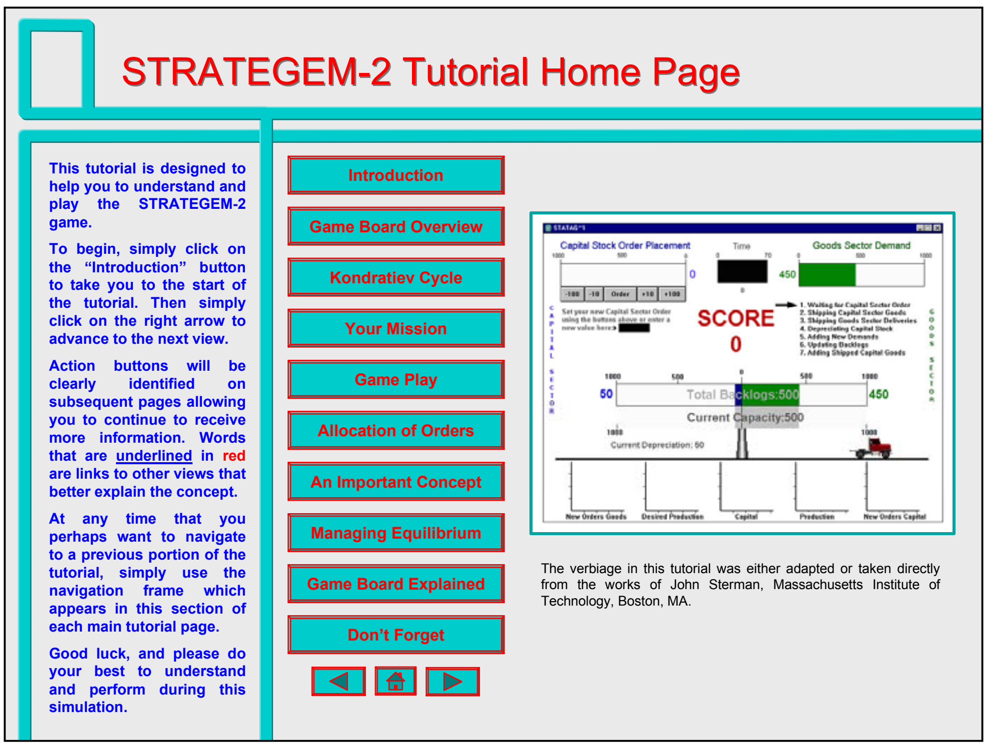
♦ These instructions were taken from the Howie STRATEGEM-2 interface (2000).

APPENDIX B

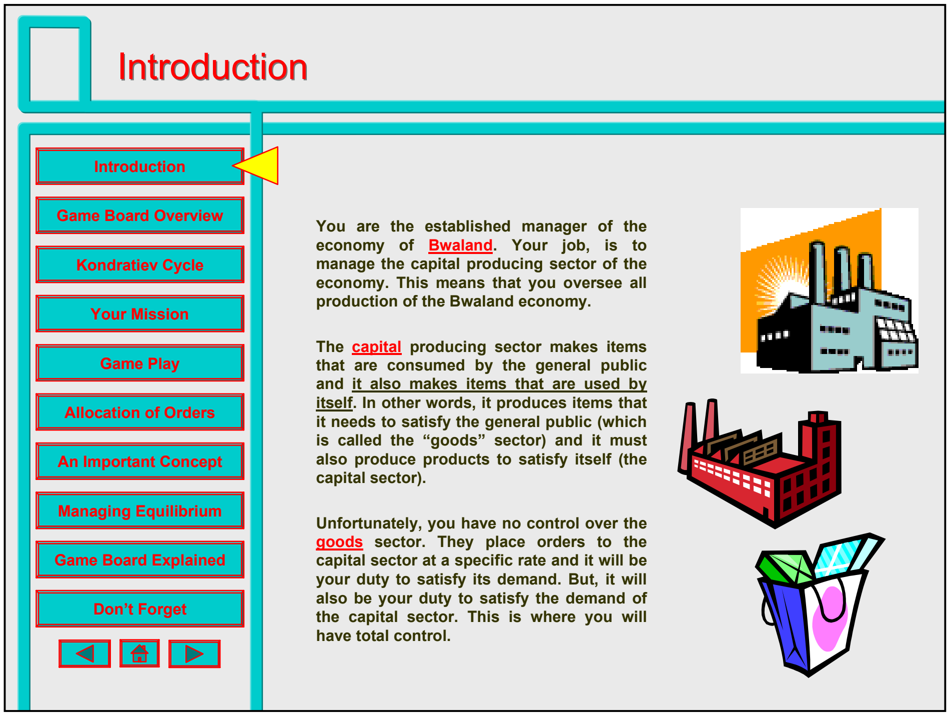
The Bois Instruction

This section will be an on-screen tutorial provided to participants. Below are pertinent views of the tutorial.

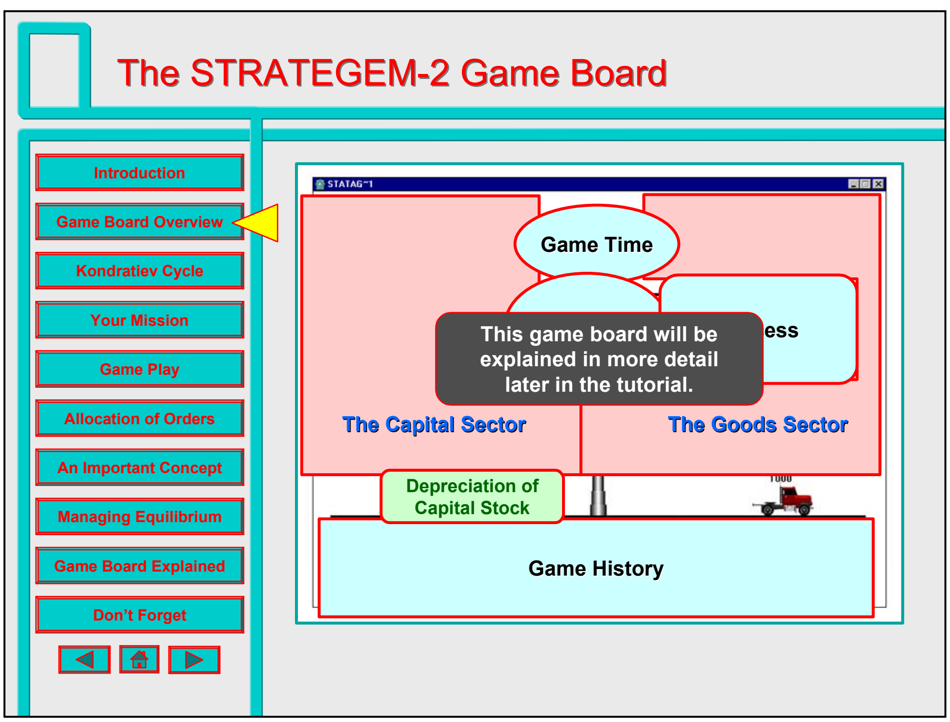
First page: Contains navigation instructions.



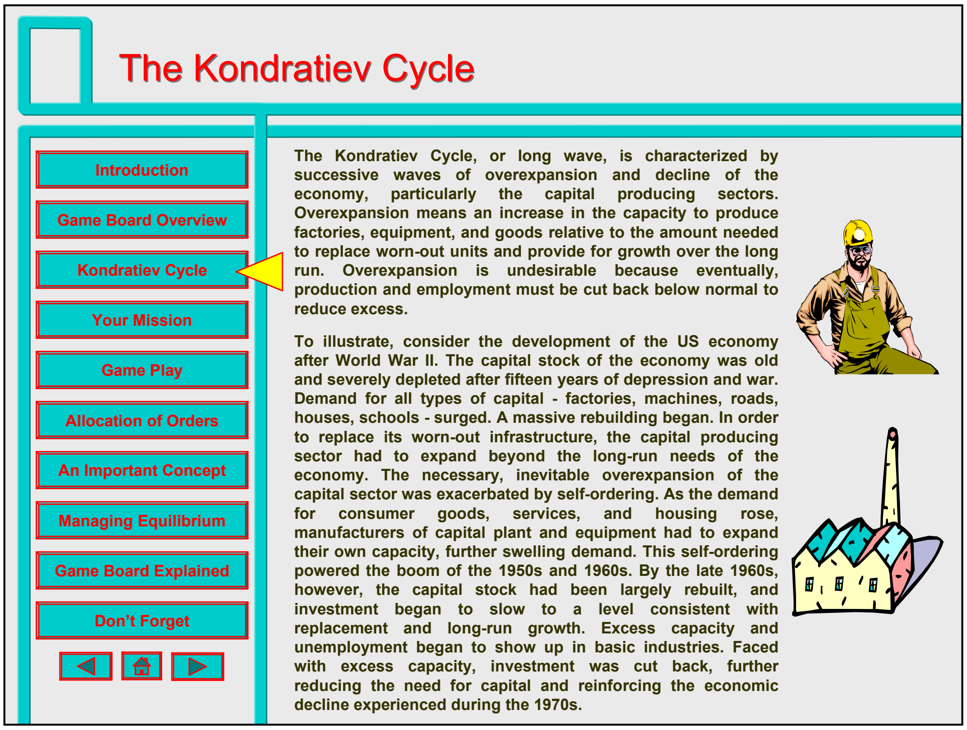
Introduction View: Contains links to other hidden pages to further define terms.



Game board overview: Has multiple overlays used to familiarize participant with game board



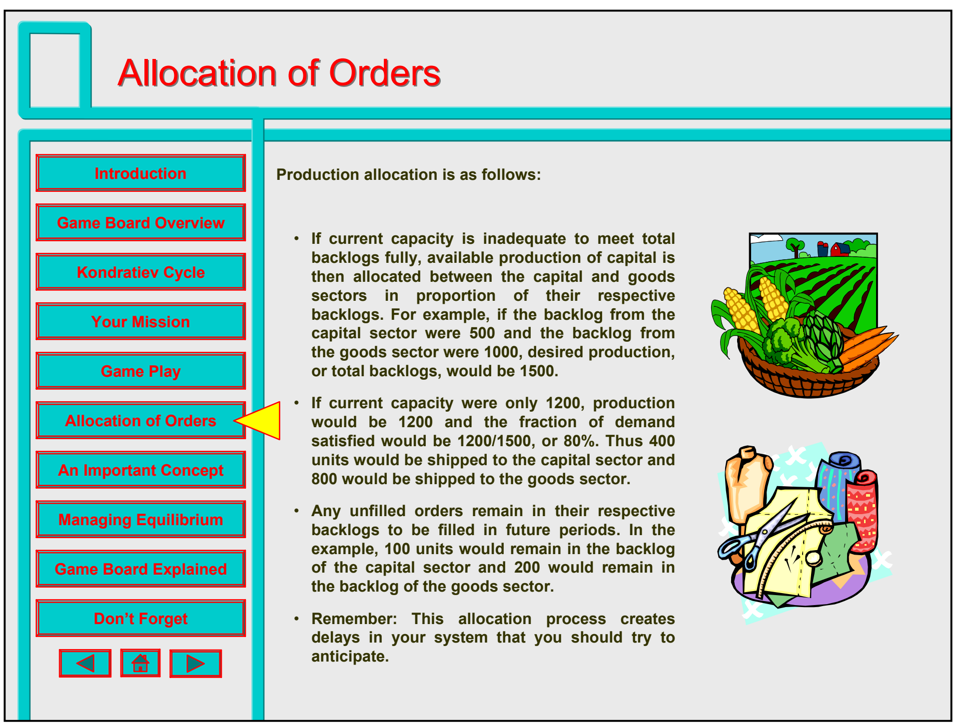
Fourth view: Explains the Kondratiev Cycle



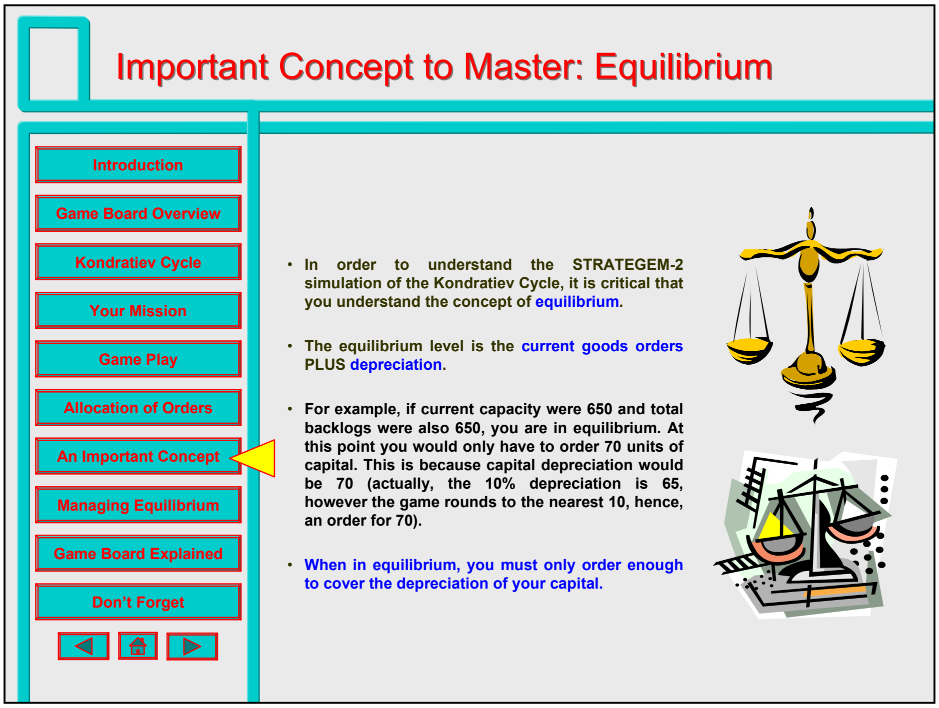
Fifth view: Explains goal / scoring of simulation with links to other explanations.

Sixth view: Outlines the play of the game.

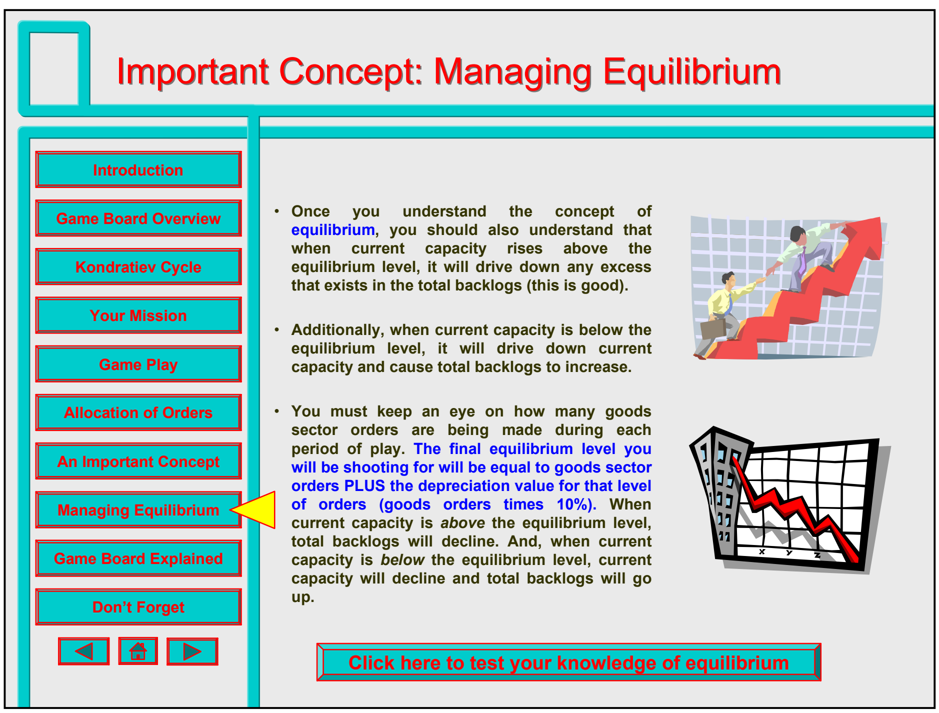
Seventh view: Explains how order allocations are made to each sector.



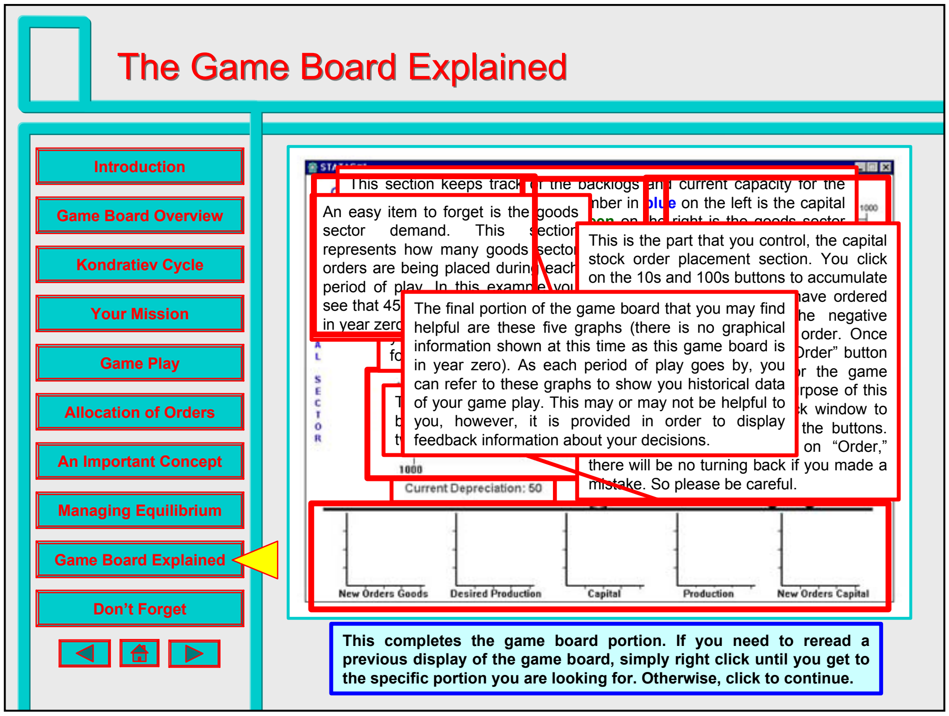
Eighth view: Introduces the concept behind game equilibrium.



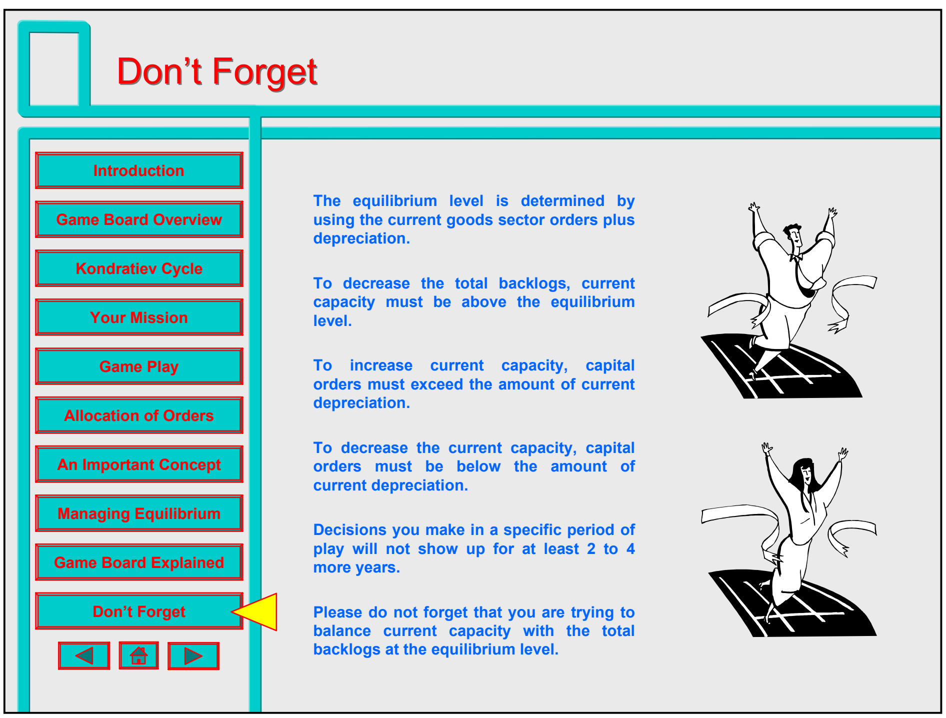
Ninth view: Explains how one manages equilibrium.
Includes link to 4-question exam used to bolster learning (not shown)



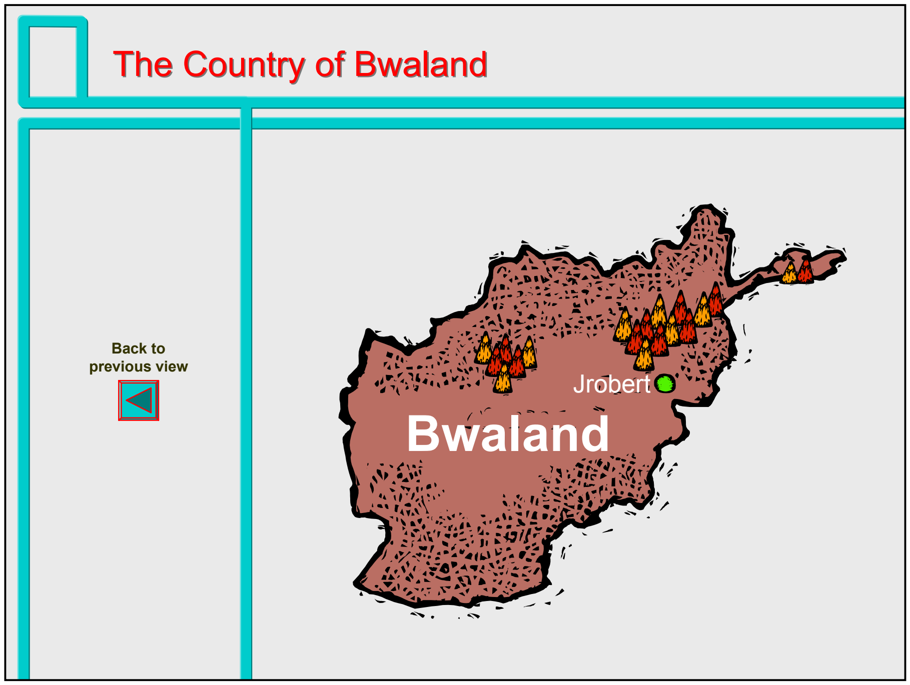
Tenth view: Explains the sections of the game board – a multiple view display.



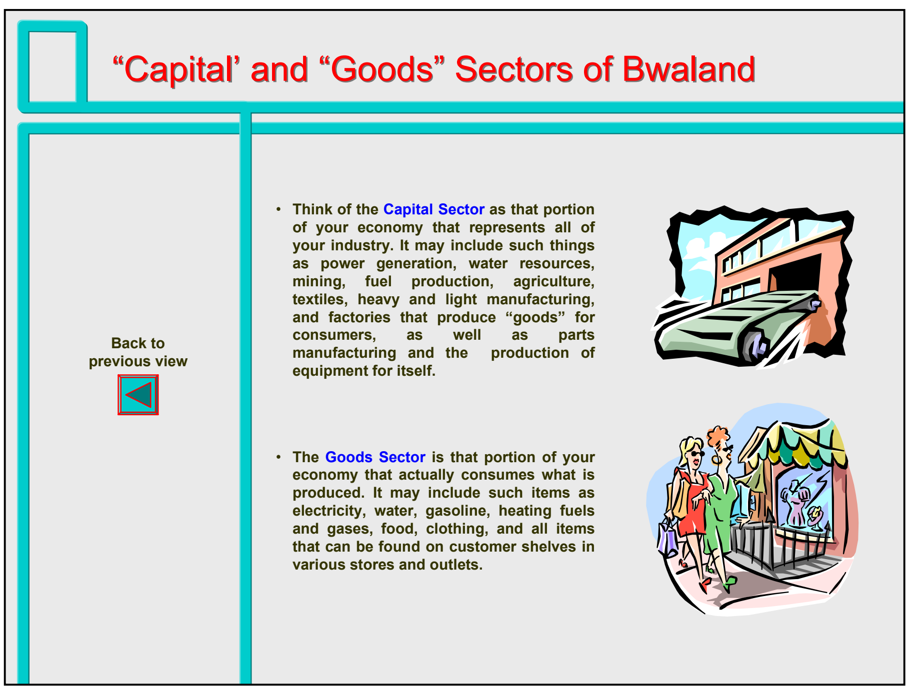
Final view: Provides tips to remember.



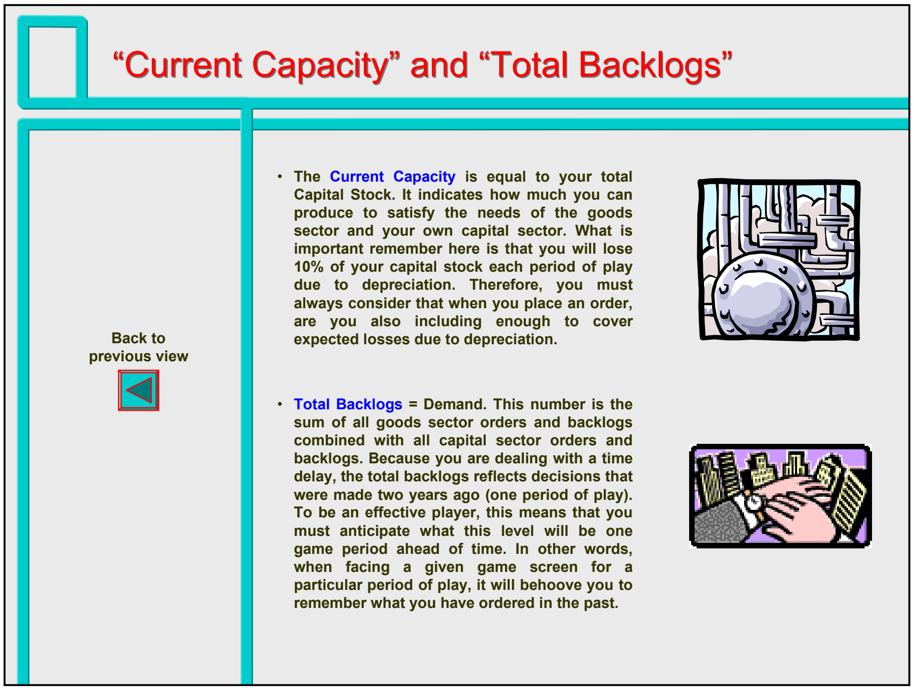
First linked view: Accessed only from another page. Used to define Bwaland.



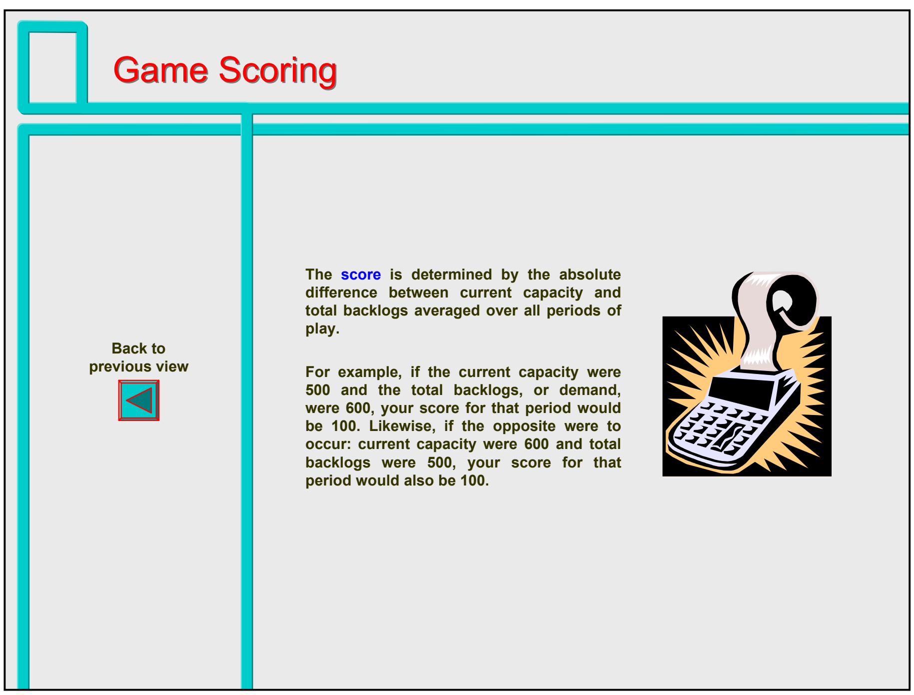
Second linked view: Used to further define the capital and goods sectors.



Third linked view: Used to further define current capacity and total backlogs.



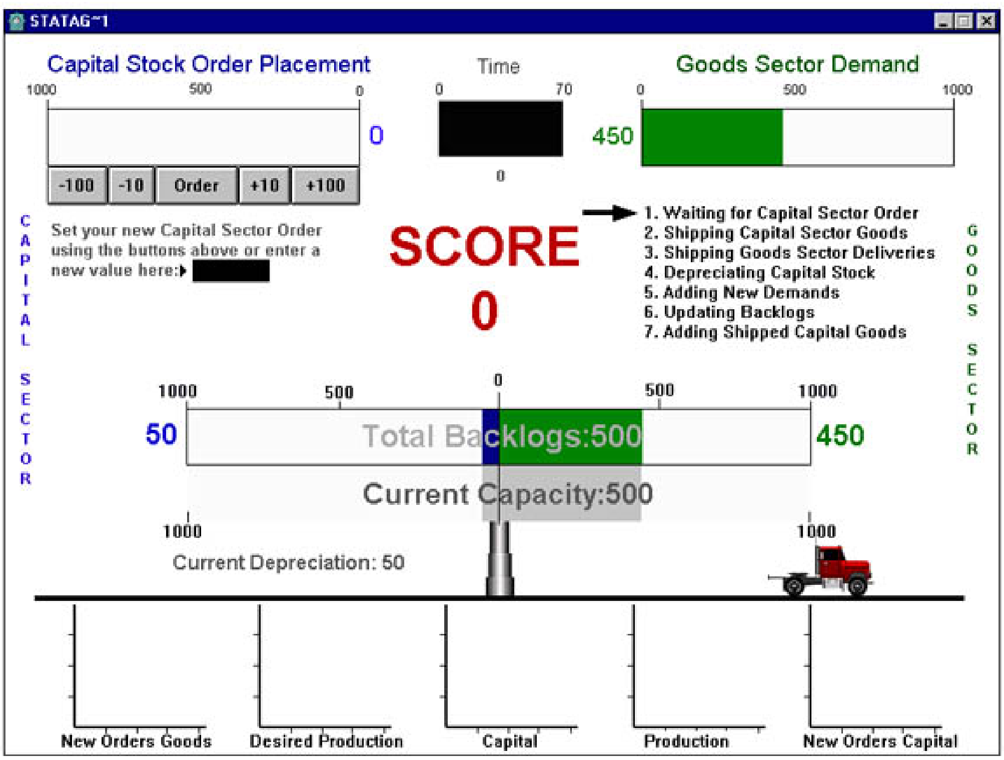
Fourth linked view: Explains how the game is scored.



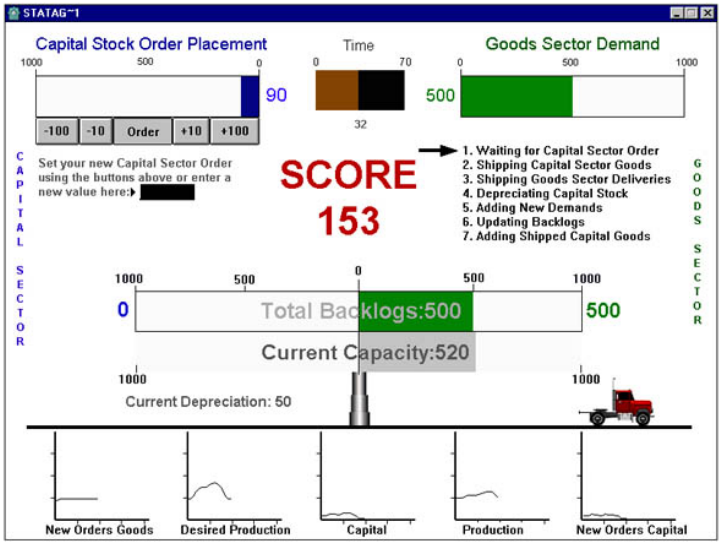
APPENDIX C

The Howie STRATEGEM-2 Interface

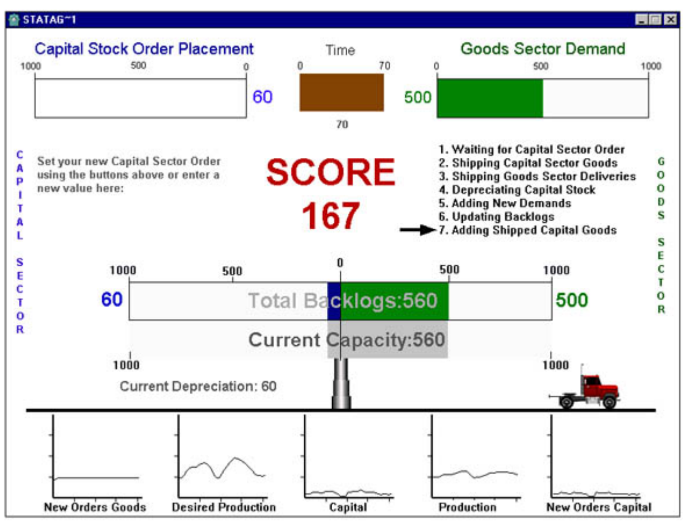
The following depiction of the STRATEGEM-2 interface shows the beginning of the game. It indicates a goods sector demand of 450 orders with an overall capacity of 500. The participant would need to only order 50 units in the capital sector (which is just enough to accommodate depreciation). The 50 capital orders combined with the 450 goods sector orders equals the 500 units of total capacity and would therefore keep the game in equilibrium and keep the score at zero.



The depiction shown below is in year 32 of a sample game played by the researcher. At this point, the researcher has allowed the current capacity (520) to be too low in order to meet the total demand, or desired production (500) plus accommodate depreciation (50). Therefore, current capacity should at least be 550 to maintain the game in equilibrium at this point. The depiction below occurred from under-ordering in the previous timeframe. In response, the researcher is ordering 90 capital units in order to boost capacity in future years.



This final depiction shows the sample game in the final year (year 70). The researcher has managed to get the game back in equilibrium (this occurred in year 68). At this point, orders required for the capital sector need only to accommodate the current depreciation (60 units). The final score for the game is 167.



APPENDIX D

Self-Assessment Survey

Please consider the following statements carefully. After each statement, circle the answer that best reflects your opinion. Would you say you strongly agree with the statement, agree, are neutral, disagree, or strongly disagree? Mark your answers accordingly on the scale for each question. As a reminder, you should answer each question as truthfully as possible. There are no wrong answers unless you are not being completely honest with yourself. Please go to the first question and begin.

For the following questions, use the following scale to rate your answer.

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

1. Regarding *this survey*, I fully understand all that is required of me from the instructions.
2. Regarding the *experiment*, I fully understood all that was required of me from the instructions.
3. I did my best in performing during the experiment.
4. The experiment took too much time to complete.
5. During the experiment, I sometimes forgot what I was supposed to do.
6. Time constraints/pressures made me hurry during my responses.
7. When provided with a set of decision cues to follow, I followed them all the time.
8. I found that the *knowledge survey* was very difficult to accomplish.
9. The required tasks were easy to understand.
10. There were times when I found myself bored with completing the tasks.
11. I am very interested in the outcomes of this research project.
12. I gave this experiment “my all.” I performed exactly in the manner prescribed in the verbal and written instructions.

Back to the Top